

09/674576 日本国特許庁
PATENT OFFICE
JAPANESE GOVERNMENT
PCT/JP 00/01304

REC'D 17 MAR 2000	
WIPO	PCT

別紙添付の書類に記載されている事項は下記の出願書類に記載されている事項と同一であることを証明する。

This is to certify that the annexed is a true copy of the following application as filed with this Office.

出願年月日
Date of Application:

1999年 3月 3日

出願番号
Application Number:

平成11年特許願第056107号

出願人
Applicant(s):

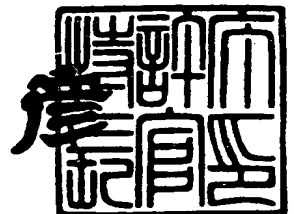
ソニー株式会社

PRIORITY DOCUMENT
SUBMITTED OR TRANSMITTED IN
COMPLIANCE WITH
RULE 17.1(a) OR (b)

2000年 1月21日

特許庁長官
Commissioner,
Patent Office

近藤 隆彦



出証番号 出証特2000-3000548

【書類名】 特許願

【整理番号】 9800978604

【提出日】 平成11年 3月 3日

【あて先】 特許庁長官殿

【国際特許分類】 G06F 15/21

【発明者】

【住所又は居所】 東京都品川区北品川6丁目7番35号 ソニー株式会社
内

【氏名】 須藤 福治

【発明者】

【住所又は居所】 東京都品川区北品川6丁目7番35号 ソニー株式会社
内

【氏名】 赤羽 誠

【発明者】

【住所又は居所】 東京都品川区北品川6丁目7番35号 ソニー株式会社
内

【氏名】 土井 利忠

【特許出願人】

【識別番号】 000002185

【氏名又は名称】 ソニー株式会社

【代表者】 出井 伸之

【代理人】

【識別番号】 100067736

【弁理士】

【氏名又は名称】 小池 晃

【選任した代理人】

【識別番号】 100086335

【弁理士】

【氏名又は名称】 田村 榮一

【選任した代理人】

【識別番号】 100096677

【弁理士】

【氏名又は名称】 伊賀 誠司

【手数料の表示】

【予納台帳番号】 019530

【納付金額】 21,000円

【提出物件の目録】

【物件名】 明細書 1

【物件名】 図面 1

【物件名】 要約書 1

【包括委任状番号】 9707387

【プルーフの要否】 要

【書類名】 明細書

【発明の名称】 コンテンツ選択システム、コンテンツ選択クライアント、コンテンツ選択サーバ及びコンテンツ選択方法

【特許請求の範囲】

【請求項 1】 サーバがネットワークを介してクライアントにコンテンツを選択させるためのコンテンツ選択用情報を送信するコンテンツ選択システムにおいて、

上記クライアントは、入力された音声情報を上記ネットワークを介して上記サーバに送信する入力情報送信手段と、上記サーバから上記ネットワークを介して上記コンテンツ選択用情報を受信して出力する出力手段とを備え、

上記サーバは、各コンテンツに関する 1 又は 2 以上の作成情報を各コンテンツごとに記憶する作成情報記憶手段と、上記クライアントから上記ネットワークを介して受信した音声情報と上記作成情報とに基づいて上記コンテンツ選択用情報を作成するコンテンツ選択用情報作成手段と、このコンテンツ選択用情報作成手段により作成された上記コンテンツ選択用情報を上記クライアントにネットワークを介して送信するコンテンツ選択用情報送信手段とを備えること

を特徴とするコンテンツ選択システム。

【請求項 2】 上記クライアントは、上記入力情報送信手段に入力された上記音声情報について音声認識を行う音声認識手段を備え、

上記入力情報送信手段は、この音声認識手段により音声認識が行われた音声認識処理済みの音声情報を上記サーバに送信し、

上記コンテンツ選択用情報作成手段は、上記クライアントから受信した上記音声認識処理済みの音声情報と上記作成情報とに基づいて、上記コンテンツ選択用情報を作成すること

を特徴とする請求項 1 記載のコンテンツ選択システム。

【請求項 3】 上記サーバは、上記クライアントからネットワークを介して受信した上記音声情報について音声認識を行う音声認識手段を備え、

上記コンテンツ選択用情報作成手段は、この音声認識手段により音声認識が行われた音声情報と上記作成情報とに基づいて、上記コンテンツ選択用情報を作成

すること

を特徴とする請求項1記載のコンテンツ選択システム。

【請求項4】 上記作成情報記憶手段は、音声情報からなる作成情報を記憶し

上記コンテンツ選択用情報作成手段は、音響的な特徴量における上記音声情報と上記作成情報記憶手段に記憶された上記各作成情報との類似性を演算し、この類似性の値が所定の閾値を越えた作成情報に対応するコンテンツの数が所定の数以上の場合には、上記音声情報とは異なる音声情報を取得して類似性の演算を繰り返し、上記類似性の値が所定の閾値を越えた上記作成情報に対応する上記コンテンツの数が所定の数未満の場合には、上記類似性の値が所定の閾値を越えた上記作成情報に対応する上記コンテンツについてのコンテンツ選択用情報を作成すること

を特徴とする請求項1記載のコンテンツ選択システム。

【請求項5】 上記コンテンツ選択用情報作成手段は、音響的な特徴量における上記音声情報と上記作成情報記憶手段に記憶された上記各作成情報との類似性を演算し、この類似性の値が所定の閾値を越えた作成情報に対応するコンテンツの数が所定の数以上の場合には、上記各作成情報の範疇ごとのばらつきを演算し、このばらつきが最も大きい範疇の作成情報に基づいてコンテンツ選択用情報を作成し、上記類似性の値が所定の閾値を越えた上記作成情報に対応する上記コンテンツの数が所定の数未満の場合には、上記類似性の値が所定の閾値を越えた上記作成情報に対応する上記コンテンツについてのコンテンツ選択用情報を作成すること

を特徴とする請求項1記載のコンテンツ選択システム。

【請求項6】 上記音声認識手段は、上記音声情報についての音声認識が正確に行われたのか否かを判定する判定手段を有し、この判定手段により正確に音声認識が行われたと判定された場合には、上記音声認識処理済みの音声情報を出力し、上記判定手段により正確に音声認識が行われなかったと判定された場合には、再度上記入力情報送信手段に入力された音声情報について音声認識を行って上記音声認識処理済みの音声情報を出力し、

上記入力情報送信手段は、上記音声認識手段により音声認識が行われた音声認識処理済みの音声情報を上記サーバに送信し、

上記コンテンツ選択用情報作成手段は、上記クライアントから受信した上記音声認識処理済みの音声情報と上記作成情報とに基づいて、上記コンテンツ選択用情報を作成すること

を特徴とする請求項 2 記載のコンテンツ選択システム。

【請求項 7】 上記音声認識手段は、上記音声情報についての音声認識が正確に行われたのか否かを判定する判定手段を有し、この判定手段により正確に音声認識が行われたと判定された場合には、上記音声認識処理済みの音声情報を出力し、上記判定手段により正確に音声認識が行われなかったと判定された場合には、再度上記クライアントからネットワークを介して受信した音声情報について音声認識を行って上記音声認識処理済みの音声情報を出力し、

上記コンテンツ選択用情報作成手段は、上記音声認識手段により音声認識が行われた音声情報と上記作成情報とに基づいて、上記コンテンツ選択用情報を作成すること

を特徴とする請求項 3 記載のコンテンツ選択システム。

【請求項 8】 サーバからネットワークを介して送信されたコンテンツを選択させるためのコンテンツ選択用情報を出力するコンテンツ選択クライアントにおいて、

入力された音声情報を上記ネットワークを介して上記サーバに送信する入力情報送信手段と、

上記サーバから上記ネットワークを介して上記コンテンツ選択用情報を受信して出力する出力手段とを備えること

を特徴とするコンテンツ選択クライアント。

【請求項 9】 上記入力情報送信手段は、入力された上記音声情報について音声認識を行う音声認識手段を備え、この音声認識手段により音声認識が行われた音声認識処理済みの音声情報を上記サーバに送信すること

を特徴とする請求項 8 記載のコンテンツ選択クライアント。

【請求項 10】 各コンテンツに関する 1 又は 2 以上の作成情報を各コンテンツごとに記憶する作成情報記憶手段と、

クライアントからネットワークを介して受信した音声情報と上記作成情報とに基づいて上記コンテンツを選択するためのコンテンツ選択用情報を作成するコンテンツ選択用情報作成手段と、

このコンテンツ選択用情報作成手段により作成された上記コンテンツ選択用情報を上記クライアントにネットワークを介して送信するコンテンツ選択用情報送信手段とを備えること

を特徴とするコンテンツ選択サーバ。

【請求項 11】 上記クライアントからネットワークを介して受信した上記音声情報について音声認識を行う音声認識手段を備え、

上記コンテンツ選択用情報作成手段は、この音声認識手段により音声認識が行われた音声情報と上記作成情報とに基づいて、上記コンテンツ選択用情報を作成すること

を特徴とする請求項 10 記載のコンテンツ選択サーバ。

【請求項 12】 サーバは、各コンテンツに関する 1 又は 2 以上の作成情報をこの各コンテンツごとに記憶し、

クライアントは、入力された音声情報をネットワークを介して上記サーバに送信し、

上記サーバは、上記クライアントから上記ネットワークを介して受信した音声情報と上記作成情報とに基づいて上記コンテンツを選択するためのコンテンツ選択用情報を作成し、この作成されたコンテンツ選択用情報を上記クライアントにネットワークを介して送信し、

上記クライアントは、上記サーバから上記ネットワークを介して上記コンテンツ選択用情報を受信し、このコンテンツ選択用情報に基づいて選択することを要求するコンテンツ選択要求情報を出力すること

を特徴とするコンテンツ選択方法。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】

本発明は、クライアントのマイク等に音声情報を入力することにより、サーバからネットワークを介してクライアントが受信して出力した複数の項目からなるコンテンツのリストの中から、どのコンテンツを取得したいのかを操作者に選択させるコンテンツ選択システム、コンテンツ選択クライアント、コンテンツ選択サーバ及びコンテンツ選択方法に関するものである。

【0002】

【従来の技術】

近年、携帯性が重視される例えば携帯電話等の携帯端末が広く用いられており、携帯性の向上のために携帯端末の軽量化及び小型化が望まれている。携帯端末は、小型化に伴って、キーボード等の入力装置を小さくする又は無くす必要がある。そこで、携帯端末においては、従来の入力装置に比べて比較的場所をとらない音声認識装置が注目を集めている。即ち、携帯端末は、音声認識装置を用いることにより端末自身を小型化できるだけでなく、利用者の利便性を更に向上させることができる。

【0003】

また、サーバ側とクライアント側との間で構築されるコンテンツ選択システムにおいては、音声等のデータ圧縮技術や放送／通信分野でのデジタル信号処理技術の発達により、所謂Music On Demand（以下、MODという。）等の音声情報を提供するサービスが実現されている。

【0004】

【発明が解決しようとする課題】

しかしながら、入力された情報を認識する手段として音声認識装置を携帯端末に用いた場合、音声認識の誤りをなくすことは非常に困難である。しかも、従来の音声認識装置における入力された音声情報を正確に認識できる確率は、あまり高くはない。

【0005】

また、MODサービスにおいて、提供してもらいたい音楽等を指定するために入力された音声情報について音声認識装置が誤認識した場合には、ユーザは、また同じ音声情報を音声認識装置に入力しなければならない。しかも、音声認識装置が次にユーザが入力する音声情報を正確に音声認識できるとは限らないので、ユーザは、何度か同じ音声情報を音声認識装置に入力しなければならないこともあり、こういう場合には非常に不便である。

【0006】

さらに、ユーザがいくつかの音声情報を音声認識装置に入力することにより提供してもらいたい音楽等を指定した場合に、この入力された音声情報の条件に当てはまる音楽等が複数存在するときがある。このようなときに、ユーザが提供されることを要求している音楽等を特定する方法は、非常に難しい。

【0007】

そこで、本発明は、このような実情に鑑みてなされたものであり、クライアントのマイク等に音声情報を入力することにより、サーバからネットワークを介してクライアントが受信して出力した複数の項目であるコンテンツの中からどのコンテンツを取得したいのかを操作者に選択させるコンテンツ選択システム、コンテンツ選択クライアント、コンテンツ選択サーバ及びコンテンツ選択方法を提供することを目的とするものである。

【0008】

【課題を解決するための手段】

上述の目的を達成するために、本発明に係るコンテンツ選択システムは、サーバがネットワークを介してクライアントにコンテンツを選択させるためのコンテンツ選択用情報を送信するコンテンツ選択システムにおいて、上記クライアントは、入力された音声情報を上記ネットワークを介して上記サーバに送信する入力情報送信手段と、上記サーバから上記ネットワークを介して上記コンテンツ選択用情報を受信して出力する出力手段とを備え、上記サーバは、各コンテンツに関する1又は2以上の作成情報を各コンテンツごとに記憶する作成情報記憶手段と、上記クライアントから上記ネットワークを介して受信した音声情報と上記作成

情報とに基づいて上記コンテンツ選択用情報を作成するコンテンツ選択用情報作成手段と、このコンテンツ選択用情報作成手段により作成された上記コンテンツ選択用情報を上記クライアントにネットワークを介して送信するコンテンツ選択用情報送信手段とを備えることを特徴とする。

【0009】

このコンテンツ選択システムでは、クライアントは、入力された音声情報をネットワークを介してサーバに送信し、サーバは、クライアントからネットワークを介して受信した音声情報と作成情報とに基づいてコンテンツ選択用情報を作成して、この作成されたコンテンツ選択用情報をクライアントにネットワークを介して送信する。

【0010】

また、本発明に係るコンテンツ選択クライアントは、サーバからネットワークを介して送信されたコンテンツを選択させるためのコンテンツ選択用情報を出力するコンテンツ選択クライアントにおいて、入力された音声情報を上記ネットワークを介して上記サーバに送信する入力情報送信手段と、上記サーバから上記ネットワークを介して上記コンテンツ選択用情報を受信して出力する出力手段とを備えることを特徴とする。

【0011】

このコンテンツ選択クライアントでは、入力された音声情報をネットワークを介してサーバに送信し、また、サーバからネットワークを介してコンテンツ選択用情報を受信して出力する。

【0012】

さらに、本発明に係るコンテンツ選択サーバは、各コンテンツに関する1又は2以上の作成情報を各コンテンツごとに記憶する作成情報記憶手段と、クライアントからネットワークを介して受信した音声情報と上記作成情報とに基づいて上記コンテンツを選択するためのコンテンツ選択用情報を作成するコンテンツ選択用情報作成手段と、このコンテンツ選択用情報作成手段により作成された上記コンテンツ選択用情報を上記クライアントにネットワークを介して送信するコンテンツ選択用情報送信手段とを備えることを特徴とする。

【0013】

このコンテンツ選択サーバでは、各コンテンツに関する1又は2以上の作成情報を各コンテンツごとに記憶し、クライアントからネットワークを介して受信した音声情報と作成情報とに基づいてコンテンツを選択するためのコンテンツ選択用情報を作成して、この作成されたコンテンツ選択用情報をクライアントにネットワークを介して送信する。

【0014】

さらにまた、本発明に係るコンテンツ選択方法は、サーバは、各コンテンツに関する1又は2以上の作成情報をこの各コンテンツごとに記憶し、クライアントは、入力された音声情報をネットワークを介して上記サーバに送信し、上記サーバは、上記クライアントから上記ネットワークを介して受信した音声情報と上記作成情報とに基づいて上記コンテンツを選択するためのコンテンツ選択用情報を作成し、この作成されたコンテンツ選択用情報を上記クライアントにネットワークを介して送信し、上記クライアントは、上記サーバから上記ネットワークを介して上記コンテンツ選択用情報を受信し、このコンテンツ選択用情報に基づいて選択することを要求するコンテンツ選択要求情報を出力することを特徴とする。

【0015】

このコンテンツ選択方法では、クライアントは、入力された音声情報をネットワークを介して上記サーバに送信し、サーバは、クライアントからネットワークを介して受信した音声情報と作成情報とに基づいてコンテンツを選択するためのコンテンツ選択用情報を作成して、この作成されたコンテンツ選択用情報をクライアントにネットワークを介して送信する。そして、クライアントは、サーバからネットワークを介してコンテンツ選択用情報を受信し、このコンテンツ選択用情報に基づいて選択することを要求するコンテンツ選択要求情報を出力する。

【0016】

【発明の実施の形態】

以下、本発明を適用した実施の形態について、図面を参照しながら説明する。

本発明を適用した実施の形態であるコンテンツ選択システムは、例えば、International Multimedia Telecommunica

tions-2000 (以下、IMT-2000という。)を用いたネットワークにおいて、所謂Music On Demand (以下、MODという。)サービスを利用するための情報提供システムの中で用いられるものである。本発明を適用した実施の形態であるコンテンツ選択システムを図1に示す。なお、IMT-2000とは、International Telecommunication Union (ITU) が、21世紀のデジタル携帯電話等に採用する通信方式をいう。

【0017】

この図1に示すように、コンテンツ選択システム1は、例えば通信機能を備えるPersonal Digital Assistant (以下、PDAという。)であるクライアント2と、各種サーバを備えるサーバシステム3と、例えばIMT-2000を介してクライアント2とサーバシステム3との間の無線回線の設定や制御を行う無線設備である公衆用基地局 (以下、基地局という。)4とを備える。

【0018】

クライアント2は、ユーザが入力した例えば音声情報を処理する入力部2aと、コンテンツとこのコンテンツを選択するためのコンテンツ選択用情報等を出力する出力部2bと、サーバシステム3から受信したコンテンツ選択用情報等を記憶するクライアント用記憶部2cと、サーバシステム3との接続を確立するための接続処理及びクライアント2が用いる通信方式に従った通信処理を行う通信部2dと、上記入力部2aと出力部2bとクライアント用記憶部2cと通信部2dの各処理部の機能を制御する制御部2eとを有する。

【0019】

ここで、コンテンツ選択用情報とは、サーバシステム3からクライアント2に送信する候補となるいくつかのコンテンツに関する情報が項目ごとに示されたりリスト等をいう。

【0020】

入力部2aは、例えばマイク等の音声入力装置により構成される。この入力部2aのマイク等は、例えば、ユーザが聴くことを希望する音楽情報であるコンテ

ンツ（以下、音楽コンテンツという。）等をサーバシステム3から送信してもらうために、ユーザが入力した音声情報を制御部2eに供給する。なお、入力部2aは、例えばキーボードやジョグダイヤル等の音声以外の入力装置により構成されていてもよい。

【0021】

出力部2bは、例えばスピーカ等の音声出力装置により構成される。この出力部2bのスピーカ等は、例えば、サーバシステム3から後述する通信部2dと制御部2eとを介して送信されてきた音楽コンテンツ等を再生する。

【0022】

なお、出力部2bは、例えば、液晶ディスプレイのような表示デバイスにより構成されていてもよい。また、出力部2bは、例えば Institute of Electrical and Electronics Engineers 1394（以下、IEEE1394という。）端子などのデジタル出力端子を備えることにより、サーバシステム3から受信した音楽コンテンツ等のコンテンツを外部端子に出力してもよい。さらに、出力部2bは、サーバシステム3から基地局4を介して受信した、例えば、歌詞、楽譜、作詞者、作曲者、編曲者、収録アルバム、発売日等のコンテンツに付随する情報を、音声情報としてスピーカ等から又は映像情報として液晶ディスプレイ等から出力してもよい。または、出力部2bは、サーバシステム3から受信した音楽コンテンツ等のコンテンツを再生中に、このコンテンツに付随する情報を液晶ディスプレイ等に表示してもよい。

【0023】

クライアント用記憶部2cは、例えば、メモリーカード、Mini Disk（以下、MDという。）、Digital Audio Taper recorder（以下、DATという。）等のクライアント2から取り外し可能、又はフラッシュメモリやハードディスク等のクライアント2から取り外し不可能な記憶装置により構成される。このクライアント用記憶部2cは、サーバシステム3から通信部2dと制御部2eとを介して送信されてきた例えば音楽コンテンツや音楽コンテンツのリスト等のコンテンツ情報を記憶する。また、クライアント用記憶

部 2 c は、各クライアント 2 が有する個別の例えば IMT-2000 用のユーザ ID 等を記憶し、制御部 2 e からの制御信号に基づいて制御部 2 e と通信部 2 d とを介してサーバシステム 3 にこのユーザ ID 等を送信する。

【0024】

通信部 2 d は、例えば、ベースバンド部や Radio Frequency (以下、RF という。) 部により構成される。この通信部 2 d は、制御部 2 e から供給された入力音声情報をサーバシステム 3 に送信したり、サーバシステム 3 からコンテンツ選択用情報等を受信し、この受信したコンテンツ選択用情報等を制御部 2 e に供給したりする。

【0025】

制御部 2 e は、例えば CPU により構成される。この制御部 2 e は、例えば、サーバシステム 3 から基地局 4 を介してコンテンツ選択用情報等を受信するための制御プログラムを備えており、この制御プログラムを実行することにより出力部 2 b、通信部 2 d 及びクライアント用記憶部 2 c の制御を行う。なお、制御部 2 e は、入力部 2 a と接続されており、この入力部 2 a から入力される音声情報等に基づいて上記制御プログラムにおける各制御を行う。

【0026】

サーバシステム 3 は、例えば、クライアント 2 から受信した入力音声情報について音声認識を行う音声認識サーバ 5 と、この音声認識サーバ 5 が認識した入力音声情報に基づいてクライアント 2 との接続認証処理等を行う認証サーバ 6 と、音声認識サーバ 5 が認識した入力音声情報と後述する作成情報とに基づいてコンテンツ選択用情報を作成するコンテンツ選択用情報作成サーバ (以下、情報作成サーバという。) 7 と、このコンテンツ選択用情報に基づいてユーザが選択したコンテンツをクライアント 2 に送信する情報提供サーバ 8 と、クライアント 2 に送信するコンテンツ等を記憶するデータベース 9 とを有する。

【0027】

音声認識サーバ 5 は、クライアント 2 の通信部 2 d から基地局 4 を介して受信した情報が音声情報である場合に、図示しない音響分析部、音響モデル及び言語モデル等に基づいて受信した音声情報の音声認識を行う。そして、音声認識サー

バ5は、この音声認識を行った結果である音声情報（以下、認識音声情報という。）を後述する認証サーバ6や選択サーバ7に供給する。なお、音声認識サーバ5は、例えば音声認識を行った結果が正確であるか否かをユーザに確認してもらうために、認識音声情報をクライアント2に送信してもよい。

【0028】

ここで、音響分析部とは、入力音声情報に対して音響的な特徴量系列の抽出を行う処理部をいう。また、音響モデルとは、この音響分析部で抽出された個々の特徴量との音響的な類似性の評価を行うための評価式を用いて、音声の部分的或いは全体的な特徴量系列のパターンとの類似性の評価を行うための情報をいう。さらに、言語モデルとは、音響モデルの接続に関する制約を与えるための情報をいう。

【0029】

認証サーバ6は、音声認識サーバ5から供給された認識音声情報等に基づいて、例えば、Remote Authentication Dial In User Service（以下、RADIUSという。）を用いたクライアント2への認証処理を行う。また、認証サーバ6は、音声認識サーバ5から供給された認識音声情報等に基づいて、例えば、Point to Point Protocol（以下、PPPという。）等を用いたクライアント2からの接続処理を行う。

【0030】

ここで、RADIUSとは、ダイヤルインのためのモデムを備えたネットワーク・アクセス・サーバに接続するというダイヤルインのユーザの認証を行うダイヤルアップ接続ユーザ認証システムをいう。また、PPPとは、公衆電話回線などでインターネットへ接続するためのプロトコルをいう。

【0031】

情報作成サーバ7は、音声認識サーバ5から供給された認識音声情報と、後述するデータベース9の作成情報記憶部9bから取得した作成情報とに基づいて、コンテンツ選択用情報を作成する。即ち、情報作成サーバ7は、後述するコンテンツ選択用情報の作成方法の中からいくつかの作成方法を用いることによりコン

テンツ選択用情報を作成する。そして、情報作成サーバ7は、作成したコンテンツ選択用情報をクライアント2に送信し、ユーザにサーバシステム3から提供してもらいたいコンテンツを入力部2aを用いることにより選択してもらう。

【0032】

情報提供サーバ8は、クライアント2が情報作成サーバ7から受信したコンテンツ選択用情報に基づいてユーザが選択した音楽コンテンツ等に関する情報をクライアント2から受信することにより、クライアント2に送信する音楽コンテンツ等をコンテンツ情報記憶部9cから取得してクライアント2に送信する。

【0033】

データベース9は、例えばユーザIDとパスワード等の属性情報を記憶する属性情報記憶部9aと、各コンテンツに関する1又は2以上の作成情報をこの各コンテンツごとに記憶する作成情報記憶部9bと、クライアント2に送信するコンテンツ等を記憶するコンテンツ情報記憶部9cとを有する。

【0034】

なお、上述したサーバシステム3の各サーバは、同一のサーバで構成してもよい。

【0035】

以上のように構成されたコンテンツ選択システム1では、サーバシステム3は、クライアント2からコンテンツ選択用情報を作成するための入力音声情報等を受信し、この入力音声情報等について音声認識を行い、音声認識を行った入力音声情報である認識音声情報と作成情報とに基づいてコンテンツ選択用情報を作成し、この作成したコンテンツ選択用情報をクライアント2に送信する。そして、クライアント2は、ユーザに聴くことを希望する音楽コンテンツ等を選択させ、この選択した音楽コンテンツ等に関する情報をサーバシステム3に送信する。サーバシステム3は、クライアント2から受信した音楽コンテンツ等に関する情報に基づいてコンテンツ情報記憶部9cからユーザが選択した音楽コンテンツ等をクライアント2に送信する。クライアント2は、この送信された音楽コンテンツ等を出力部2bのスピーカ等を用いて再生する。

【0036】

つぎに、コンテンツ選択システム1を用いた情報提供システムにおいて、サーバシステム3がクライアント2に音楽コンテンツ等を提供するMODサービスが開始されてから終了するまでの一連の処理について、具体的な音楽コンテンツ等の選択方法を用いた一例を、図2に示すフローチャートに従って説明する。

【0037】

この図2に示すフローチャートでは、例えば、IMT-2000用の端末であるクライアント2は、IMT-2000のデータ回線を用いた基地局4を介して、サーバシステム3に例えば電話をかける等して予め接続要求を出している。その後、クライアント2は、サーバシステム3との接続が確立する。即ち、この図2に示すフローチャートでは、クライアント2とサーバシステム3との接続が確立している状態から説明する。

【0038】

まず、図2のステップS1において、ユーザがMODサービスを利用するためのサービス開始要求情報を、クライアント2に備えられた入力部2aのマイク等に音声で入力した場合に、制御部2eは、入力部2aに入力された音声情報であるサービス開始要求情報（以下、サービス開始要求音声情報という。）を、この入力部2aから供給される。制御部2eは、入力部2aから供給されたサービス開始要求音声情報を通信部2dに供給する。通信部2dは、供給されたサービス開始要求音声情報を基地局4を介してサーバシステム3の音声認識サーバ5に送信する。

【0039】

音声認識サーバ5は、通信部2dから基地局4を介して受信したサービス開始要求音声情報について、図示しない音響分析部、音響モデル及び言語モデル等に基づき音声認識を行う。音声認識サーバ5は、この音声認識を行った結果であるサービス開始要求音声情報（以下、認識サービス開始要求音声情報という。）を認証サーバ6に供給する。

【0040】

認証サーバ6は、音声認識サーバ5から認識サービス開始要求音声情報が供給

されると、ユーザが入力部 2 a に音声等でユーザ ID とパスワードを入力することを要求するユーザ認証要求情報をクライアント 2 に送信する。ここで、このユーザ ID とパスワードとは、クライアント 2 を所有するユーザが MOD サービスを利用する権利を有するか否かを認証するための情報をいう。

【0041】

制御部 2 e は、認証サーバ 6 から通信部 2 d を介して受信したユーザ認証要求情報を、出力部 2 b のスピーカ等に音声として出力させる。

【0042】

続いて、ステップ S 2 において、ユーザが出力部 2 b のスピーカ等から音声出力されたユーザ認証要求情報に従って、入力部 2 a のマイク等にユーザ ID とパスワードを音声で入力した場合に、制御部 2 e は、入力部 2 a に入力された音声情報であるユーザ ID とパスワードを、この入力部 2 a から供給される。制御部 2 e は、入力部 2 a から供給されたユーザ ID とパスワードを通信部 2 d に供給する。通信部 2 d は、供給されたユーザ ID とパスワードを基地局 4 を介してサーバシステム 3 の音声認識サーバ 5 に送信する。

【0043】

音声認識サーバ 5 は、通信部 2 d から基地局 4 を介して受信したユーザ ID とパスワードについて音声認識を行う。音声認識サーバ 5 は、この音声認識を行った結果であるユーザ ID とパスワード（以下、認識ユーザ ID ・パスワードという。）を認証サーバ 6 に供給する。

【0044】

認証サーバ 6 は、データベース 9 の属性情報記憶部 9 a からユーザ ID とパスワードを取得し、この取得したユーザ ID とパスワードに基づいて音声認識サーバ 5 から供給された認識ユーザ ID ・パスワードの認証処理を行う。

【0045】

認証サーバ 6 は、この認証ができない場合には、サービス開始不許可情報をクライアント 2 に送信する。制御部 2 e は、認証サーバ 6 から通信部 2 d を介して受信したサービス開始不許可情報に基づいて、MOD サービスが開始されない旨を、出力部 2 b のスピーカ等に音声として出力させる。ユーザがこの MOD サー

ビスが開始されない旨をスピーカ等から聞いて、改めて入力部 2 a のマイク等にユーザ ID とパスワードを正確に音声で入力した場合には、制御部 2 e は、入力部 2 a に再度入力された音声情報であるユーザ ID とパスワードを、この入力部 2 a から供給される。

【0046】

一方、認証サーバ 6 は、上述の認証ができた場合には、認証ができたことを知らせる認証成功情報を情報作成サーバ 7 に供給する。情報作成サーバ 7 は、この認証成功情報を認証サーバ 6 から供給された場合には、サービス開始許可情報と、ユーザに曲のタイトルを入力するように要求するタイトル入力要求情報とをクライアント 2 に送信する。制御部 2 e は、情報作成サーバ 7 から通信部 2 d を介して受信したサービス開始許可情報とタイトル入力要求情報とを、出力部 2 b のスピーカ等に音声として出力させる。そして、処理は、ステップ S 3 へ進む。

【0047】

続いて、ステップ S 3 において、ユーザがサービス開始許可情報とタイトル入力要求情報とをスピーカ等から聞いた後に、制御部 2 e は、ユーザが入力部 2 a のマイク等にサーバシステム 3 から提供されることを希望する曲のタイトル（以下、タイトルという。）を音声で入力した音声情報（以下、タイトル音声情報という。）を、この入力部 2 a から供給される。制御部 2 e は、入力部 2 a から供給されたタイトル音声情報を通信部 2 d に供給する。通信部 2 d は、供給されたタイトル音声情報を基地局 4 を介してサーバシステム 3 の音声認識サーバ 5 に送信する。

【0048】

音声認識サーバ 5 は、通信部 2 d から基地局 4 を介して受信したタイトル音声情報について音声認識を行う。音声認識サーバ 5 は、この音声認識を行った結果であるタイトル音声情報（以下、認識タイトル音声情報という。）を情報作成サーバ 7 に供給する。

【0049】

続いて、ステップ S 4 において、情報作成サーバ 7 は、音声認識サーバ 5 から供給された認識タイトル音声情報と、データベース 9 の作成情報記憶部 9 b から

取得した 1 又は 2 以上の各タイトルに関する作成情報（以下、タイトル作成情報という。）とに基づいて、サーバシステム 3 から提供されることを希望する曲に関するコンテンツ選択用情報（以下、曲コンテンツ選択用情報という。）を作成する。

【 0 0 5 0 】

具体的には、情報作成サーバ 7 は、音響的な特徴量における、認識タイトル音声情報と、作成情報記憶部 9 b に記憶されている情報提供システムで提供可能な曲のタイトルに関する各タイトル作成情報との類似性を、図示しない作成情報演算処理部を用いて演算する。この情報作成サーバ 7 は、作成情報記憶部 9 b に記憶されている全ての各タイトル作成情報に対して、音響的な特徴量における認識タイトル音声情報との類似性を演算する。情報作成サーバ 7 は、この演算値が予め設定された所定の閾値である X（以下、閾値 X という。）を越えたものを、ユーザがサーバシステム 3 から提供されることを希望する曲の候補になると判断する。

【 0 0 5 1 】

そして、演算値が閾値 X を越えたタイトル作成情報に対応する曲の数が所定の定数である N（以下、定数 N という。）個未満である場合には、処理は、ステップ S 8 へ進む。

【 0 0 5 2 】

一方、演算値が閾値 X を越えたタイトル作成情報に対応する曲の数が定数 N 個以上の場合には、情報作成サーバ 7 は、候補となる曲の数を絞り込むために、ユーザに曲のパフォーマの名前を入力するように要求するパフォーマ名入力要求情報をクライアント 2 に送信する。制御部 2 e は、情報作成サーバ 7 から通信部 2 d を介して受信したパフォーマ名入力要求情報を、出力部 2 b のスピーカ等に音声として出力させる。そして、処理は、ステップ S 5 へ進む。

【 0 0 5 3 】

続いて、ステップ S 5 において、ユーザがパフォーマ名入力要求情報をスピーカ等から聞いた後に、制御部 2 e は、ユーザが入力部 2 a のマイク等にサーバシステム 3 から提供されることを希望する曲のパフォーマの名前（以下、パフォー

マ名という。)を音声で入力した音声情報(以下、パフォーマ名音声情報という。)を、この入力部2aから供給される。制御部2eは、入力部2aから供給されたパフォーマ名音声情報を通信部2dに供給する。通信部2dは、供給されたパフォーマ名音声情報を基地局4を介してサーバシステム3の音声認識サーバ5に送信する。

【0054】

音声認識サーバ5は、通信部2dから基地局4を介して受信したパフォーマ名音声情報について音声認識を行う。音声認識サーバ5は、この音声認識を行った結果であるパフォーマ名音声情報(以下、認識パフォーマ名音声情報という。)を情報作成サーバ7に供給する。

【0055】

続いて、ステップS6において、情報作成サーバ7は、音声認識サーバ5から供給された認識パフォーマ名音声情報と、データベース9の作成情報記憶部9bから取得した1又は2以上の各パフォーマ名に関する作成情報(以下、パフォーマ名作成情報という。)とに基づいて、曲コンテンツ選択用情報を作成する。但し、この場合のパフォーマ名作成情報とは、タイトルによる絞り込みで演算値が閾値Xを越えたものの中から取得したパフォーマ名に関する作成情報のみをいう。従って、作成情報記憶部9bは、予め、コンテンツである各曲に関するタイトル、パフォーマ名、ジャンル等を記憶している。

【0056】

具体的には、情報作成サーバ7は、音響的な特徴量における、認識パフォーマ名音声情報と、作成情報記憶部9bに記憶されている情報提供システムで提供可能な曲のパフォーマ名に関する各パフォーマ名作成情報との類似性を、図示しない作成情報演算処理部を用いて演算する。この情報作成サーバ7は、作成情報記憶部9bに記憶されている全ての各パフォーマ名作成情報に対して、音響的な特徴量における認識パフォーマ名音声情報との類似性を演算する。情報作成サーバ7は、この演算値が閾値Xを越えたものを、ユーザがサーバシステム3から提供されることを希望する曲の候補になると判断する。

【0057】

そして、演算値が閾値Xを越えたパフォーマ名作成情報に対応する曲の数が定数N個未満である場合、即ち、タイトル及びパフォーマ名の両方を用いた場合における演算値が閾値Xを越えたタイトル作成情報及びパフォーマ名作成情報に対応する曲の数が定数N個未満であるときには、処理は、ステップS8へ進む。

【0058】

一方、タイトル及びパフォーマ名の両方を用いた場合における演算値が閾値Xを越えたパフォーマ名作成情報に対応する曲の数が定数N個以上のときには、情報作成サーバ7は、候補となる曲の数を絞り込むために、ユーザに曲のジャンルを入力するように要求するジャンル入力要求情報をクライアント2に送信する。制御部2eは、情報作成サーバ7から通信部2dを介して受信したジャンル入力要求情報を、出力部2bのスピーカ等に音声として出力させる。そして、処理は、ステップS7へ進む。

【0059】

続いて、ステップS7において、ユーザがジャンル入力要求情報をスピーカ等から聞いた後に、制御部2eは、ユーザが入力部2aのマイク等にサーバシステム3から提供されることを希望する曲のジャンル（以下、ジャンルという。）を音声で入力した音声情報（以下、ジャンル音声情報という。）を、この入力部2aから供給される。制御部2eは、入力部2aから供給されたジャンル音声情報を通信部2dに供給する。通信部2dは、供給されたジャンル音声情報を基地局4を介してサーバシステム3の音声認識サーバ5に送信する。

【0060】

音声認識サーバ5は、通信部2dから基地局4を介して受信したジャンル音声情報について音声認識を行う。音声認識サーバ5は、この音声認識を行った結果であるジャンル音声情報（以下、認識ジャンル音声情報という。）を情報作成サーバ7に供給する。

【0061】

情報作成サーバ7は、音声認識サーバ5から供給された認識ジャンル音声情報と、データベース9の作成情報記憶部9bから取得した1又は2以上の各ジャン

ルに関する作成情報（以下、ジャンル作成情報という。）とに基づいて、曲コンテンツ選択用情報を作成する。但し、この場合のジャンル作成情報とは、タイトル及びパフォーマ名による絞り込みで演算値が閾値Xを越えたものの中から取得したジャンルに関する作成情報のみをいう。

【0062】

具体的には、情報作成サーバ7は、音響的な特徴量における、認識ジャンル音声情報と、作成情報記憶部9bに記憶されている情報提供システムで提供可能な曲のジャンルに関する各ジャンル作成情報との類似性を、図示しない作成情報演算処理部を用いて演算する。この情報作成サーバ7は、作成情報記憶部9bに記憶されている全ての各ジャンル作成情報に対して、音響的な特徴量における認識ジャンル音声情報との類似性を演算する。情報作成サーバ7は、この演算値が閾値Xを越えたものを、ユーザがサーバシステム3から提供されることを希望する曲の候補になると判断する。

【0063】

そして、演算値が閾値Xを越えたジャンル作成情報に対応する曲の数が定数N個以下である場合、即ち、タイトル、パフォーマ名及びジャンルを用いた場合における演算値が閾値Xを越えたタイトル作成情報、パフォーマ名作成情報及びジャンル作成情報に対応する曲の数が定数N個未満であるときには、処理は、ステップS8へ進む。

【0064】

一方、タイトル、パフォーマ名及びジャンルを用いた場合における演算値が閾値Xを越えたタイトル作成情報、パフォーマ名作成情報及びジャンル作成情報に対応する曲の数が定数N個以上のときには、情報作成サーバ7は、演算値がジャンル音声情報に最も類似していたものから順にN個をクライアント2に送信する。つまり、この場合も、処理は、ステップS8へ進む。

【0065】

続いて、ステップS8において、演算値が閾値Xを越えた上記作成情報に対応する曲の数が定数N個以上の場合に、情報作成サーバ7が曲コンテンツ選択用情報をクライアント2に送信したときには、制御部2eは、情報作成サーバ7から

通信部 2 d を介して受信した演算値がジャンル音声情報に最も類似していたものからの順番である N 個の曲コンテンツ選択用情報を、出力部 2 b のスピーカ等に音声として出力させる。

【0066】

この場合、制御部 2 e は、例えば、タイトルに関する演算値とパフォーマ名に関する演算値とジャンルに関する演算値の平均等を利用して、入力された音声情報により類似した曲から順番に並べられた N 個の曲のリスト等である曲コンテンツ選択用情報を、出力部 2 b のスピーカ等に音声として出力させる。

【0067】

一方、演算値が閾値 X を越えた上記作成情報に対応する曲の数が定数 N 個未満である場合に、情報作成サーバ 7 が曲コンテンツ選択用情報をクライアント 2 に送信したときには、制御部 2 e は、情報作成サーバ 7 から通信部 2 d を介して受信した N 個以下の曲コンテンツ選択用情報を、出力部 2 b のスピーカ等に音声として出力させる。

【0068】

この場合、制御部 2 e は、例えば、タイトルに関する演算値とパフォーマ名に関する演算値とジャンルに関する演算値の平均等を利用して、入力された音声情報により類似した曲から順番に並べられた曲のリスト等である曲コンテンツ選択用情報を、出力部 2 b のスピーカ等に音声として出力させる。

【0069】

なお、この曲コンテンツ選択用情報とは、例えば、曲の名前、曲の一部、パフォーマ名、タイアップしているドラマの名前、その曲が放送された番組の名前、曲の製作者の名前など、ユーザが入力した音声情報に対する結果となり得るもの全ての情報をいう。また、制御部 2 e は、この曲コンテンツ選択用情報を、例えば、液晶ディスプレイ等を用いて、テキスト情報、画像情報等として出力部 2 b に出力させてもよい。

【0070】

続いて、ステップ S 9 において、ユーザが曲コンテンツ選択用情報をスピーカ等から聞いた後に、制御部 2 e は、この曲コンテンツ選択用情報である曲のリス

トの中からユーザが選択した曲に関する情報を入力部 2 a のマイク等に音声で入力された音声情報を、この入力部 2 a から供給される。

【0071】

具体的には、制御部 2 e は、例えば、曲コンテンツ選択用情報である曲のリストの項目に割り振られた番号を入力部 2 a に音声で入力された音声情報を、この入力部 2 a から供給される。

【0072】

また、制御部 2 e は、例えば、曲コンテンツ選択用情報である曲の名前を入力部 2 a に音声で入力された音声情報を、この入力部 2 a から供給される。

【0073】

さらに、制御部 2 e は、例えば、出力部 2 b が曲コンテンツ選択用情報である曲のリストを順番に音声で出力している場合で、現在、音声で出力されている曲をユーザが選択したいときに、「この曲」や「OK」等という言葉が入力部 2 a に音声で入力された音声情報を、この入力部 2 a から供給される。この場合、入力部 2 a は、キー入力装置やジョグダイヤル等により構成されていてもよい。

【0074】

さらにまた、制御部 2 e は、例えば、出力部 2 b が曲コンテンツ選択用情報である曲の一部分を順番に再生している場合で、現在、再生している曲をユーザが選択したいときに、「この曲」や「OK」等という言葉が入力部 2 a に音声で入力された音声情報を、この入力部 2 a から供給される。この場合も、入力部 2 a は、キー入力装置やジョグダイヤル等により構成されていてもよい。

【0075】

さらにまた、制御部 2 e は、例えば、予めクライアント 2 がサーバシステム 3 からユーザの趣向にあった曲のリストをダウンロードしてある場合に、その曲のリストの中からユーザが選択した曲の名前が入力部 2 a に音声で入力された音声情報を、この入力部 2 a から供給される。このユーザの趣向にあった曲のリストは、ユーザが予め設定しておいてもよい。

【0076】

以上の様な操作をされることにより、曲コンテンツ選択用情報である曲のリス

ト等の中から、ユーザがサーバシステム3から提供されることを希望する曲が選択される。

【0077】

続いて、ステップS10において、制御部2eは、入力部2aに入力されたユーザが曲のリストの中から選択した曲に関する音声情報（以下、選択音声情報という。）を、この入力部2aから供給される。制御部2eは、入力部2aから供給された選択音声情報を通信部2dに供給する。通信部2dは、供給された選択音声情報を基地局4を介してサーバシステム3の音声認識サーバ5に送信する。

【0078】

音声認識サーバ5は、通信部2dから基地局4を介して受信した選択音声情報について音声認識を行う。音声認識サーバ5は、この音声認識を行った結果である選択音声情報（以下、認識選択音声情報という。）を、クライアント2に送信する。

【0079】

制御部2eは、音声認識サーバ5から通信部2dを介して受信した認識選択音声情報を、出力部2bのスピーカ等に音声として出力させる。ユーザは、この認識選択音声情報をスピーカ等から聞くことにより、自分が選択した曲をサーバシステム3が正しく認識したかどうか確認する。

【0080】

続いて、ステップS11において、情報提供サーバ8は、音声認識サーバ5から供給された認識選択音声情報に基づき、クライアント2に送信する音楽等のコンテンツをコンテンツ情報記憶部9cから取得してクライアント2に送信する。

【0081】

制御部2eは、情報提供サーバ8から通信部2dを介して受信した音楽等のコンテンツを、出力部2bのスピーカ等に再生させる。

【0082】

続いて、ステップS12において、出力部2bからのコンテンツの再生が終了し、ユーザが引き続きMODサービスを利用する場合には、処理は、ステップS3へ戻る。

【0083】

一方、ユーザがMODサービスの利用を終了するためのサービス終了要求情報を、入力部2aのマイク等に音声で入力した場合には、制御部2eは、入力部2aに入力された音声情報であるサービス終了要求情報（以下、サービス終了要求音声情報という。）を、この入力部2aから供給される。制御部2eは、入力部2aから供給されたサービス終了要求音声情報を通信部2dに供給する。通信部2dは、供給されたサービス終了要求音声情報を基地局4を介してサーバシステム3の音声認識サーバ5に送信する。

【0084】

音声認識サーバ5は、通信部2dから基地局4を介して受信したサービス終了要求音声情報について、図示しない音響分析部、音響モデル及び言語モデル等に基づき音声認識を行う。音声認識サーバ5は、この音声認識を行った結果であるサービス終了要求音声情報（以下、認識サービス終了要求音声情報という。）を情報作成サーバ7に供給する。

【0085】

情報作成サーバ7は、音声認識サーバ5から認識サービス終了要求音声情報が供給されると、サービス終了許可情報をクライアント2に送信する。

【0086】

制御部2eは、情報作成サーバ7から通信部2dを介して受信したサービス終了許可情報を、制御部2eのスピーカ等に音声として出力させる。

【0087】

ユーザは、このサービス終了許可情報をスピーカ等から音声として聞くことによって、MODサービスが終了したことを認識する。

【0088】

以上のように処理されることにより、ユーザは、サーバシステム3からMODサービスを受けることができる。

【0089】

なお、上述した図2に示すフローチャートを用いた一連の処理では、入力部2aは、音声で情報が入力されるマイク等で構成されているが、例えばキー入力さ

れるキーボード等で構成されていてもよい。同様に、出力部 2 b は、音声出力するスピーカ等で構成されているが、例えば、映像出力する液晶ディスプレイ等で構成されていてもよい。また、出力部 2 b は、これらのスピーカや液晶ディスプレイ等を用いて、映像及び音声を同時に出力してもよい。

【0090】

また、図 2 に示すフローチャートを用いた一連の処理では、音声認識サーバ 5 は、認識音声情報を直接認証サーバ 6 や情報作成サーバ 7 に供給しているが、例えば音声認識を行った結果が正確であるか否かをユーザに確認してもらうために、認識音声情報をクライアント 2 に送信してもよい。

【0091】

そして、出力部 2 b が音声認識サーバ 5 から受信してスピーカ等で音声出力した認識音声情報が間違っていたとユーザが認識した場合に、制御部 2 e は、例えば、「取り直し」や「やり直し」等の音声情報をユーザが入力部 2 a に再度正確に入力し直した音声情報を、この入力部 2 a から供給されてもよい。勿論、制御部 2 e は、例えばユーザが「クリアキー」を押すことにより、入力部 2 a に再度正確に情報等を入力し直した音声情報を、この入力部 2 a から供給されてもよい。

【0092】

さらに、図 2 に示すフローチャートを用いた一連の処理では、ユーザがサービス開始要求情報を入力部 2 a に音声で入力することによりサービスが開始されているが、クライアント 2 がサーバシステム 3 に例えば電話をかける等して接続されたらサービスが開始されるとしてもよい。

【0093】

さらにまた、図 2 に示すフローチャートを用いた一連の処理では、コンテンツ選択システム 1 を用いた情報提供システムにおいて MOD サービスを利用している途中でこのサービスの利用を終了する場合、制御部 2 e は、ネットワークとの接続の切断を命令する切断音声情報をユーザが入力部 2 a に入力し直した音声情報を、この入力部 2 a から供給されることにより、基地局 4 を介したサーバシステム 3 との接続を切断して MOD サービスを終了させてもよい。勿論、制御部 2

eは、例えばユーザが入力部2aの図示しない「END-KEY」等を押すことにより、基地局4を介したサーバシステム3との接続を切断してMODサービスを終了させてもよい。

【0094】

さらにまた、図2に示すフローチャートのステップS2において、制御部2eは、ユーザが入力部2aに音声で入力したユーザIDとパスワードとを、この入力部2aから供給されてもよい。しかし、例えばIMT-2000用のクライアント2の制御部2eは、ユーザにより音声で入力部2aに入力されたパスワードのみを、この入力部2aから供給されてもよい。IMT-2000用の各クライアント2は、全て他のクライアント2と違うユーザIDを予め保持しているからである。

【0095】

つぎに、上述した曲コンテンツ選択用情報の作成方法以外の作成方法を利用したコンテンツ選択システム1を用いた情報提供システムにおいて、MODサービスが開始されてから終了するまでの一連の処理について、図3に示すフローチャートに従って説明する。この処理は、上述した処理と異なり、次にどういう情報が入力されればコンテンツ選択用情報を速く作成できるのかをサーバシステム3側が判断し、その判断された情報をユーザに入力させるものである。

【0096】

この図3に示すフローチャートでは、例えば、IMT-2000用の端末であるクライアント2は、IMT-2000のデータ回線を用いた基地局4を介して、サーバシステム3に例えば電話をかける等して予め接続要求を出している。その後、クライアント2は、サーバシステム3との接続が確立する。即ち、この図3に示すフローチャートでは、クライアント2とサーバシステム3との接続が確立している状態から説明する。

【0097】

なお、この図3に示す第2の実施の形態における処理は、図2に示すフローチャートにおけるステップS4からステップS8までの処理を、図3に示すフローチャートにおけるステップS24からステップS33までの処理に換えたもので

ある。

【0098】

まず、図3のステップS21において、ユーザがサービス開始要求情報を、クライアント2に備えられた入力部2aのマイク等に音声で入力した場合に、制御部2eは、入力部2aに入力されたサービス開始要求音声情報を、この入力部2aから供給される。制御部2eは、入力部2aから供給されたサービス開始要求音声情報を通信部2dに供給する。通信部2dは、供給されたサービス開始要求音声情報を基地局4を介してサーバシステム3の音声認識サーバ5に送信する。

【0099】

音声認識サーバ5は、通信部2dから基地局4を介して受信したサービス開始要求音声情報について、図示しない音響分析部、音響モデル及び言語モデル等に基づき音声認識を行う。音声認識サーバ5は、この音声認識を行った結果である認識サービス開始要求音声情報を認証サーバ6に供給する。

【0100】

認証サーバ6は、音声認識サーバ5から認識サービス開始要求音声情報が供給されると、クライアント2を所有するユーザがMODサービスを利用する権利を有するか否かを認証するためのユーザIDとパスワードを、ユーザが入力部2aに音声等で入力することを要求するユーザ認証要求情報を、クライアント2に送信する。

【0101】

制御部2eは、認証サーバ6から通信部2dを介して受信したユーザ認証要求情報を、出力部2bのスピーカ等に音声として出力させる。

【0102】

続いて、ステップS22において、ユーザが出力部2bのスピーカ等から音声出力されたユーザ認証要求情報に従って、入力部2aのマイク等にユーザIDとパスワードを音声で入力した場合に、制御部2eは、入力部2aに入力された音声情報であるユーザIDとパスワードを、この入力部2aから供給される。制御部2eは、入力部2aから供給されたユーザIDとパスワードを通信部2dに供給する。通信部2dは、供給されたユーザIDとパスワードを基地局4を介して

サーバシステム 3 の音声認識サーバ 5 に送信する。

【0103】

音声認識サーバ 5 は、通信部 2 d から基地局 4 を介して受信したユーザ ID とパスワードについて音声認識を行う。音声認識サーバ 5 は、この音声認識を行った結果である認識ユーザ ID ・パスワードを認証サーバ 6 に供給する。

【0104】

認証サーバ 6 は、データベース 9 の属性情報記憶部 9 a からユーザ ID とパスワードを取得し、この取得したユーザ ID とパスワードに基づいて音声認識サーバ 5 から供給された認識ユーザ ID ・パスワードの認証処理を行う。

【0105】

認証サーバ 6 は、この認証ができない場合には、サービス開始不許可情報をクライアント 2 に送信する。制御部 2 e は、認証サーバ 6 から通信部 2 d を介して受信したサービス開始不許可情報に基づいて、MOD サービスが開始されない旨を、出力部 2 b のスピーカ等に音声として出力させる。ユーザがこの MOD サービスが開始されない旨をスピーカ等から聞いて、改めて入力部 2 a のマイク等にユーザ ID とパスワードを正確に音声で入力した場合には、制御部 2 e は、入力部 2 a に再度入力された音声情報であるユーザ ID とパスワードを、この入力部 2 a から供給される。

【0106】

一方、認証サーバ 6 は、上述の認証ができた場合には、認証ができたことを知らせる認証成功情報を情報作成サーバ 7 に供給する。情報作成サーバ 7 は、この認証成功情報を認証サーバ 6 から供給された場合には、サービス開始許可情報と、ユーザに曲のタイトルを入力するように要求するタイトル入力要求情報とをクライアント 2 に送信する。制御部 2 e は、情報作成サーバ 7 から通信部 2 d を介して受信したサービス開始許可情報とタイトル入力要求情報とを、出力部 2 b のスピーカ等に音声として出力させる。そして、処理は、ステップ S 2 3 へ進む。

【0107】

続いて、ステップ S 2 3 において、ユーザがサービス開始許可情報とタイトル入力要求情報とをスピーカ等から聞いた後に、制御部 2 e は、ユーザが入力部 2

aのマイク等に音声で入力したタイトル音声情報を、この入力部2aから供給される。制御部2eは、入力部2aから供給されたタイトル音声情報を通信部2dに供給する。通信部2dは、供給されたタイトル音声情報を基地局4を介してサーバシステム3の音声認識サーバ5に送信する。

【0108】

音声認識サーバ5は、通信部2dから基地局4を介して受信したタイトル音声情報について音声認識を行う。音声認識サーバ5は、この音声認識を行った結果である認識タイトル音声情報を情報作成サーバ7に供給する。

【0109】

続いて、ステップS24において、情報作成サーバ7は、音声認識サーバ5から供給された認識タイトル音声情報と、データベース9の作成情報記憶部9bから取得した1又は2以上の各タイトル作成情報とに基づいて、曲コンテンツ選択用情報を作成する。

【0110】

具体的には、情報作成サーバ7は、音響的な特徴量における、認識タイトル音声情報と、作成情報記憶部9bに記憶されている情報提供システムで提供可能な曲のタイトルに関する各タイトル作成情報との類似性を、図示しない作成情報演算処理部を用いて演算する。この情報作成サーバ7は、作成情報記憶部9bに記憶されている全ての各タイトル作成情報に対して、音響的な特徴量における認識タイトル音声情報との類似性を演算する。情報作成サーバ7は、この演算値が閾値Xを越えたものを、ユーザがサーバシステム3から提供されることを希望する曲の候補になると判断する。

【0111】

そして、演算値が閾値Xを越えたタイトル作成情報に対応する曲の数が定数N個未満である場合には、処理は、ステップS33へ進む。

【0112】

ステップS25において、演算値が閾値Xを越えたタイトル作成情報に対応する曲の数が定数N個以上の場合には、情報作成サーバ7は、候補となる曲の数を速く絞り込むために、作成情報記憶部9bからタイトルの入力情報から決定され

た候補に関するパフォーマ名及びジャンルの音声としての作成情報を取得して用いることにより、このパフォーマ名及びジャンルの各作成情報の範疇ごとのばらつきを演算する。なお、候補となる曲の数の絞り込みは、ユーザからの音声情報の入力に対して、このばらつきの程度が大きければ大きいほど速くできる場合がある。

【0113】

続いて、ステップS26において、ジャンルの音声情報の方がばらつきの程度が大きい場合には、情報作成サーバ7は、ジャンル入力要求情報をクライアント2に送信する。制御部2eは、情報作成サーバ7から通信部2dと制御部2eとを介して受信したジャンル入力要求情報を、出力部2bのスピーカ等に音声として出力させる。そして、処理は、ステップS30へ進む。

【0114】

一方、パフォーマ名の音声情報の方がばらつきの程度が大きい場合には、情報作成サーバ7は、パフォーマ名入力要求情報をクライアント2に送信する。制御部2eは、情報作成サーバ7から通信部2dを介して受信したパフォーマ名入力要求情報を、出力部2bのスピーカ等に音声として出力させる。そして、処理は、ステップS27へ進む。

【0115】

続いて、ステップS27において、ユーザがパフォーマ名入力要求情報をスピーカ等から聞いた後に、制御部2eは、入力部2aのマイク等にパフォーマ名をユーザにより音声で入力された音声情報（以下、パフォーマ名音声情報という。）を、この入力部2aから供給される。制御部2eは、入力部2aから供給されたパフォーマ名音声情報を通信部2dに供給する。通信部2dは、供給されたパフォーマ名音声情報を基地局4を介してサーバシステム3の音声認識サーバ5に送信する。

【0116】

音声認識サーバ5は、通信部2dから基地局4を介して受信したパフォーマ名音声情報について音声認識を行う。音声認識サーバ5は、この音声認識を行った結果である認識パフォーマ名音声情報を情報作成サーバ7に供給する。

【0117】

続いて、ステップS28において、情報作成サーバ7は、音声認識サーバ5から供給された認識パフォーマ名音声情報と、データベース9の作成情報記憶部9bから取得した1又は2以上の各パフォーマ名作成情報とに基づいて、曲コンテンツ選択用情報を作成する。但し、この場合のパフォーマ名作成情報とは、タイトルによる絞り込みで演算値が閾値Xを越えたものの中から取得したパフォーマ名に関する作成情報のみをいう。従って、作成情報記憶部9bは、予め、コンテンツである各曲に関するタイトル、パフォーマ名、ジャンル等を記憶している。

【0118】

具体的には、情報作成サーバ7は、音響的な特徴量における、認識パフォーマ名音声情報と、作成情報記憶部9bに記憶されている情報提供システムで提供可能な曲のパフォーマ名に関する各パフォーマ名作成情報との類似性を、図示しない作成情報演算処理部を用いて演算する。この情報作成サーバ7は、作成情報記憶部9bに記憶されている全ての各パフォーマ名作成情報に対して、音響的な特徴量における認識パフォーマ名音声情報との類似性を演算する。情報作成サーバ7は、この演算値が閾値Xを越えたものを、ユーザがサーバシステム3から提供されることを希望する曲の候補になると判断する。

【0119】

そして、演算値が閾値Xを越えたパフォーマ名作成情報に対応する曲の数が定数N個未満である場合、即ち、タイトル及びパフォーマ名の両方を用いた場合における演算値が閾値Xを越えたタイトル作成情報及びパフォーマ名作成情報に対応する曲の数が定数N個以下であるときには、処理は、ステップS33へ進む。

【0120】

一方、タイトル及びパフォーマ名の両方を用いた場合における演算値が閾値Xを越えたタイトル作成情報及びパフォーマ名作成情報に対応する曲の数が定数N個以上のときには、情報作成サーバ7は、候補となる曲の数を絞り込むために、ユーザに曲のジャンルを入力するように要求するジャンル入力要求情報をクライアント2に送信する。制御部2eは、情報作成サーバ7から通信部2dを介して受信したジャンル入力要求情報を、出力部2bのスピーカ等に音声として出力さ

せる。そして、処理は、ステップ S 29 へ進む。

【0121】

続いて、ステップ S 29 において、ユーザがジャンル入力要求情報をスピーカ等から聞いた後に、制御部 2 e は、入力部 2 a のマイク等にジャンルをユーザにより音声で入力されたジャンル音声情報を、この入力部 2 a から供給される。制御部 2 e は、入力部 2 a から供給されたジャンル音声情報を通信部 2 d に供給する。通信部 2 d は、供給されたジャンル音声情報を基地局 4 を介してサーバシステム 3 の音声認識サーバ 5 に送信する。

【0122】

音声認識サーバ 5 は、通信部 2 d から基地局 4 を介して受信したジャンル音声情報について音声認識を行う。音声認識サーバ 5 は、この音声認識を行った結果である認識ジャンル音声情報を情報作成サーバ 7 に供給する。

【0123】

情報作成サーバ 7 は、音声認識サーバ 5 から供給された認識ジャンル音声情報と、データベース 9 の作成情報記憶部 9 b から取得した 1 又は 2 以上の各ジャンル作成情報とに基づいて、曲コンテンツ選択用情報を作成する。但し、この場合のジャンル作成情報とは、タイトル及びパフォーマ名による絞り込みで演算値が閾値 X を越えたものの中から取得したジャンルに関する作成情報のみをいう。

【0124】

具体的には、情報作成サーバ 7 は、音響的な特徴量における、認識ジャンル音声情報と、作成情報記憶部 9 b に記憶されている情報提供システムで提供可能な曲のジャンルに関する各ジャンル作成情報との類似性を、図示しない作成情報演算処理部を用いて演算する。この情報作成サーバ 7 は、作成情報記憶部 9 b に記憶されている全ての各ジャンル作成情報に対して、音響的な特徴量における認識ジャンル音声情報との類似性を演算する。情報作成サーバ 7 は、この演算値が閾値 X を越えたものを、ユーザがサーバシステム 3 から提供されることを希望する曲の候補になると判断する。

【0125】

そして、演算値が閾値 X を越えたジャンル作成情報に対応する曲の数が定数 N

個未満である場合、即ち、タイトル、パフォーマ名及びジャンルを用いた場合における演算値が閾値Xを越えたタイトル作成情報、パフォーマ名及びジャンル作成情報に対応する曲の数が定数N個未満であるときには、処理は、ステップS33へ進む。

【0126】

一方、タイトル、パフォーマ名及びジャンルを用いた場合における演算値が閾値Xを越えたタイトル作成情報、パフォーマ名及びジャンル作成情報に対応する曲の数が定数N個以上のときには、情報作成サーバ7は、演算値がジャンル音声情報に最も類似していたものから順にN個をクライアント2に送信する。つまり、この場合も、処理は、ステップS33へ進む。

【0127】

また、ステップS30からステップS32においても、上述したステップS27からステップS29と同様な処理が行われる。但し、これらのステップS30からステップS32では、ジャンル、パフォーマ名の順で音声入力される。

【0128】

続いて、ステップS33において、演算値が閾値Xを越えたジャンル作成情報に対応する曲の数が定数N個未満である場合に、情報作成サーバ7が曲コンテンツ選択用情報をクライアント2に送信したときには、制御部2eは、情報作成サーバ7から通信部2dを介して受信したN個未満の曲コンテンツ選択用情報を、出力部2bのスピーカ等に音声として出力させる。

【0129】

この場合、制御部2eは、例えば、タイトルに関する演算値とパフォーマ名に関する演算値とジャンルに関する演算値の平均等を利用して、入力された音声情報により類似した曲から順番に並べられた曲のリスト等である曲コンテンツ選択用情報を、出力部2bのスピーカ等に音声として出力させる。

【0130】

なお、この曲コンテンツ選択用情報とは、例えば、曲の名前、曲の一部、パフォーマ名、タイアップしているドラマの名前、その曲が放送された番組の名前、曲の製作者の名前など、ユーザが入力した音声情報に対する結果となり得るもの

全ての情報をいう。また、制御部 2 e は、この曲コンテンツ選択用情報を、例えば、液晶ディスプレイ等を用いて、テキスト情報、画像情報等として出力部 2 b に出力させてもよい。

【0131】

一方、演算値が閾値 X を越えたジャンル作成情報に対応する曲の数が定数 N 個以上の場合に、情報作成サーバ 7 が曲コンテンツ選択用情報をクライアント 2 に送信したときには、制御部 2 e は、情報作成サーバ 7 から通信部 2 d を介して受信した演算値がジャンル音声情報に最も近似していたものからの順番である N 個の曲コンテンツ選択用情報を、出力部 2 b のスピーカ等に音声として出力させる。

【0132】

この場合、制御部 2 e は、例えば、タイトルに関する演算値とパフォーマ名に関する演算値とジャンルに関する演算値の平均等を利用して、入力された音声情報により類似した曲から順番に並べられた N 個の曲のリスト等である曲コンテンツ選択用情報を、出力部 2 b のスピーカ等に音声として出力させる。

【0133】

続いて、ステップ S 3 4 において、ユーザが曲コンテンツ選択用情報をスピーカ等から聞いた後に、制御部 2 e は、この曲コンテンツ選択用情報である曲のリストの中からユーザが選択した曲に関する情報を入力部 2 a のマイク等に音声で入力された音声情報を、この入力部 2 a から供給される。

【0134】

具体的には、制御部 2 e は、例えば、曲コンテンツ選択用情報である曲のリストの項目に割り振られた番号を入力部 2 a に音声で入力された音声情報を、この入力部 2 a から供給される。

【0135】

また、制御部 2 e は、例えば、曲コンテンツ選択用情報である曲の名前を入力部 2 a に音声で入力された音声情報を、この入力部 2 a から供給される。

【0136】

さらに、制御部 2 e は、例えば、出力部 2 b が曲コンテンツ選択用情報である

曲のリストを順番に音声で出力している場合で、現在、音声で出力されている曲をユーザが選択したいときに、「この曲」や「OK」等という言葉が入力部 2 a に音声で入力された音声情報を、この入力部 2 a から供給される。この場合、入力部 2 a は、キー入力装置やジョグダイヤル等により構成されていてもよい。

【0137】

さらにまた、制御部 2 e は、例えば、出力部 2 b が曲コンテンツ選択用情報である曲の一部分を順番に再生している場合で、現在、再生している曲をユーザが選択したいときに、「この曲」や「OK」等という言葉が入力部 2 a に音声で入力された音声情報を、この入力部 2 a から供給される。この場合も、入力部 2 a は、キー入力装置やジョグダイヤル等により構成されていてもよい。

【0138】

さらにまた、制御部 2 e は、例えば、予めクライアント 2 がサーバシステム 3 からユーザの趣向にあった曲のリストをダウンロードしてある場合に、その曲のリストの中からユーザが選択した曲の名前が入力部 2 a に音声で入力された音声情報を、この入力部 2 a から供給される。このユーザの趣向にあった曲のリストは、ユーザが予め設定しておいてもよい。

【0139】

以上の様な操作をされることにより、曲コンテンツ選択用情報である曲のリスト等の中から、ユーザがサーバシステム 3 から提供されることを希望する曲が選択される。

【0140】

続いて、ステップ S 3 5 において、制御部 2 e は、入力部 2 a に入力された選択音声情報を、この入力部 2 a から供給される。制御部 2 e は、入力部 2 a から供給された選択音声情報を通信部 2 d に供給する。通信部 2 d は、供給された選択音声情報を基地局 4 を介してサーバシステム 3 の音声認識サーバ 5 に送信する。

【0141】

音声認識サーバ 5 は、通信部 2 d から基地局 4 を介して受信した選択音声情報について音声認識を行う。音声認識サーバ 5 は、この音声認識を行った結果であ

る認識選択音声情報をクライアント2に送信する。

【0142】

制御部2eは、音声認識サーバ5から通信部2dを介して受信した認識選択音声情報を、出力部2bのスピーカ等に音声として出力させる。ユーザは、この認識選択音声情報をスピーカ等から聞くことにより、自分が選択した曲をサーバシステム3が正しく認識したかどうか確認する。

【0143】

続いて、ステップS36において、情報提供サーバ8は、音声認識サーバ5から供給された認識選択音声情報に基づき、クライアント2に送信する音楽等のコンテンツをコンテンツ情報記憶部9cから取得してクライアント2に送信する。

【0144】

制御部2eは、情報提供サーバ8から通信部2dを介して受信した音楽等のコンテンツを、出力部2bのスピーカ等に再生させる。

【0145】

続いて、ステップS37において、出力部2bからのコンテンツの再生が終了し、ユーザが引き続きMODサービスを利用する場合には、処理は、ステップS23へ戻る。

【0146】

一方、ユーザがMODサービスの利用を終了するためのサービス終了要求情報を、入力部2aのマイク等に音声で入力した場合には、制御部2eは、入力部2aに入力されたサービス終了要求音声情報を、この入力部2aから供給される。制御部2eは、入力部2aから供給されたサービス終了要求音声情報を通信部2dに供給する。通信部2dは、供給されたサービス終了要求音声情報を基地局4を介してサーバシステム3の音声認識サーバ5に送信する。

【0147】

音声認識サーバ5は、通信部2dから基地局4を介して受信したサービス終了要求音声情報について、図示しない音響分析部、音響モデル及び言語モデル等に基づき音声認識を行う。音声認識サーバ5は、この音声認識を行った結果である認識サービス終了要求音声情報を情報作成サーバ7に供給する。

【0148】

情報作成サーバ7は、音声認識サーバ5から認識サービス終了要求音声情報が供給されると、サービス終了許可情報をクライアント2に送信する。

【0149】

制御部2eは、情報作成サーバ7から通信部2dを介して受信したサービス終了許可情報を、制御部2eのスピーカ等に音声として出力させる。

【0150】

ユーザは、このサービス終了許可情報をスピーカ等から音声として聞くことによって、MODサービスが終了したことを認識する。

【0151】

以上のように処理されることにより、次にどのような情報が入力されればコンテンツ選択用情報を速く作成できるのかをサーバシステム3側が判断するため、上述した第1の実施の形態におけるコンテンツ選択システムよりも、ユーザは、サーバシステム3からMODサービスをより速く、より正確に受けることができる。

【0152】

つぎに、上述した候補となる曲のリストを出力する方法以外の出力方法を利用したコンテンツ選択システム1を用いた情報提供システムにおいて、MODサービスが開始されてから修了するまでの一連の処理について、図4に示すフローチャートに従って説明する。この処理は、上述した処理と異なり、どのように曲コンテンツ選択用情報である曲のリストが出力されればユーザが容易に目的とする曲を選択できるのかをサーバシステム3側が判断し、その判断に基づいて曲のリスト等を出力するものである。

【0153】

この図4に示すフローチャートでは、例えば、IMT-2000用の端末であるクライアント2は、IMT-2000のデータ回線を用いた基地局4を介して、サーバシステム3に例えば電話をかける等して予め接続要求を出している。その後、クライアント2は、サーバシステム3との接続が確立する。即ち、この図4に示すフローチャートでは、クライアント2とサーバシステム3との接続が確立している状態から説明する。

【0154】

なお、この図4に示す第3の実施の形態における処理は、図2に示すフローチャートにおけるステップS4からステップS8までの処理を、図4に示すフローチャートにおけるステップS44からステップS55の処理に換えたものである。

【0155】

まず、図4のステップS41において、ユーザがサービス開始要求情報を、クライアント2に備えられた入力部2aのマイク等に音声で入力した場合に、制御部2eは、入力部2aに入力されたサービス開始要求音声情報を、この入力部2aから供給される。制御部2eは、入力部2aから供給されたサービス開始要求音声情報を通信部2dに供給する。通信部2dは、供給されたサービス開始要求音声情報を基地局4を介してサーバシステム3の音声認識サーバ5に送信する。

【0156】

音声認識サーバ5は、通信部2dから基地局4を介して受信したサービス開始要求音声情報について、図示しない音響分析部、音響モデル及び言語モデル等に基づき音声認識を行う。音声認識サーバ5は、この音声認識を行った結果である認識サービス開始要求音声情報を認証サーバ6に供給する。

【0157】

認証サーバ6は、音声認識サーバ5から認識サービス開始要求音声情報が供給されると、クライアント2を所有するユーザがMODサービスを利用する権利を有するか否かを認証するためのユーザIDとパスワードを、ユーザが入力部2aに音声等で入力することを要求するユーザ認証要求情報を、クライアント2に送信する。

【0158】

制御部2eは、認証サーバ6から通信部2dを介して受信したユーザ認証要求情報を、出力部2bのスピーカ等に音声として出力させる。

【0159】

続いて、ステップS42において、ユーザが出力部2bのスピーカ等から音声出力されたユーザ認証要求情報に従って、入力部2aのマイク等にユーザIDと

パスワードを音声で入力した場合に、制御部 2 e は、入力部 2 a に入力された音声情報であるユーザ ID とパスワードを、この入力部 2 a から供給される。制御部 2 e は、入力部 2 a から供給されたユーザ ID とパスワードを通信部 2 d に供給する。通信部 2 d は、供給されたユーザ ID とパスワードを基地局 4 を介してサーバシステム 3 の音声認識サーバ 5 に送信する。

【0160】

音声認識サーバ 5 は、通信部 2 d から基地局 4 を介して受信したユーザ ID とパスワードについて音声認識を行う。音声認識サーバ 5 は、この音声認識を行った結果である認識ユーザ ID ・パスワードを認証サーバ 6 に供給する。

【0161】

認証サーバ 6 は、データベース 9 の属性情報記憶部 9 a からユーザ ID とパスワードを取得し、この取得したユーザ ID とパスワードに基づいて音声認識サーバ 5 から供給された認識ユーザ ID ・パスワードの認証処理を行う。

【0162】

認証サーバ 6 は、この認証ができない場合には、サービス開始不許可情報をクライアント 2 に送信する。制御部 2 e は、認証サーバ 6 から通信部 2 d を介して受信したサービス開始不許可情報に基づいて、MOD サービスが開始されない旨を、出力部 2 b のスピーカ等に音声として出力させる。ユーザがこの MOD サービスが開始されない旨をスピーカ等から聞いて、改めて入力部 2 a のマイク等にユーザ ID とパスワードを正確に音声で入力した場合には、制御部 2 e は、入力部 2 a に再度入力された音声情報であるユーザ ID とパスワードを、この入力部 2 a から供給される。

【0163】

一方、認証サーバ 6 は、上述の認証ができた場合には、認証ができたことを知らせる認証成功情報を情報作成サーバ 7 に供給する。情報作成サーバ 7 は、この認証成功情報を認証サーバ 6 から供給された場合には、サービス開始許可情報と、ユーザに曲のタイトルを入力するように要求するタイトル入力要求情報とをクライアント 2 に送信する。制御部 2 e は、情報作成サーバ 7 から通信部 2 d を介して受信したサービス開始許可情報とタイトル入力要求情報とを、出力部 2 b の

スピーカ等に音声として出力させる。そして、処理は、ステップ S 4 3 へ進む。

【0164】

続いて、ステップ S 4 3 において、ユーザがサービス開始許可情報とタイトル入力要求情報とをスピーカ等から聞いた後に、制御部 2 e は、ユーザが入力部 2 a のマイク等に音声で入力したタイトル音声情報を、この入力部 2 a から供給される。制御部 2 e は、入力部 2 a から供給されたタイトル音声情報を通信部 2 d に供給する。通信部 2 d は、供給されたタイトル音声情報を基地局 4 を介してサーバシステム 3 の音声認識サーバ 5 に送信する。

【0165】

音声認識サーバ 5 は、通信部 2 d から基地局 4 を介して受信したタイトル音声情報について音声認識を行う。音声認識サーバ 5 は、この音声認識を行った結果である認識タイトル音声情報を情報作成サーバ 7 に供給する。

【0166】

続いて、ステップ S 4 4 において、情報作成サーバ 7 は、音響的な特徴量における、認識タイトル音声情報と 1 又は 2 以上の各タイトル作成情報との類似性を、図示しない作成情報演算処理部を用いて演算する。この情報作成サーバ 7 は、作成情報記憶部 9 b に記憶されている全ての各タイトル作成情報に対して、音響的な特徴量における認識タイトル音声情報との類似性を演算する。情報作成サーバ 7 は、この演算値が閾値 X を越えたものを、ユーザがサーバシステム 3 から提供されることを希望する曲の候補になると判断する。

【0167】

そして、演算値が閾値 X を越えたタイトル作成情報に対応する曲の数が定数 N 個未満である場合には、情報作成サーバ 7 が候補となる曲のリスト等である曲コンテンツ選択用情報を作成するため、処理は、ステップ S 4 8 へ進む。

【0168】

一方、演算値が閾値 X を越えたタイトル作成情報に対応する曲の数が定数 N 個以上の場合には、情報作成サーバ 7 は、候補となる曲の数を絞り込むために、ユーザに曲のパフォーマの名前を入力するように要求するパフォーマ名入力要求情報をクライアント 2 に送信する。制御部 2 e は、情報作成サーバ 7 から通信部 2

d を介して受信したパフォーマ名入力要求情報を、出力部 2 b のスピーカ等に音声として出力させる。そして、処理は、ステップ S 4 5 へ進む。

【0169】

続いて、ステップ S 4 5 において、ユーザがパフォーマ名入力要求情報をスピーカ等から聞いた後に、制御部 2 e は、ユーザが入力部 2 a のマイク等に音声で入力したパフォーマ名音声情報を、この入力部 2 a から供給される。制御部 2 e は、入力部 2 a から供給されたパフォーマ名音声情報を通信部 2 d に供給する。通信部 2 d は、供給されたパフォーマ名音声情報を基地局 4 を介してサーバシステム 3 の音声認識サーバ 5 に送信する。

【0170】

音声認識サーバ 5 は、通信部 2 d から基地局 4 を介して受信したパフォーマ名音声情報について音声認識を行う。音声認識サーバ 5 は、この音声認識を行った結果である認識パフォーマ名音声情報を情報作成サーバ 7 に供給する。

【0171】

続いて、ステップ S 4 6 において、情報作成サーバ 7 は、音響的な特徴量における、認識パフォーマ名音声情報と、作成情報記憶部 9 b に記憶されている情報提供システムで提供可能な曲のパフォーマ名に関する 1 又は 2 以上の各パフォーマ名作成情報との類似性を、図示しない作成情報演算処理部を用いて演算する。この情報作成サーバ 7 は、作成情報記憶部 9 b に記憶されている全ての各パフォーマ名作成情報に対して、音響的な特徴量における認識パフォーマ名音声情報との類似性を演算する。情報作成サーバ 7 は、この演算値が閾値 X を越えたものを、ユーザがサーバシステム 3 から提供されることを希望する曲の候補になると判断する。

【0172】

そして、演算値が閾値 X を越えたパフォーマ名作成情報に対応する曲の数が定数 N 個未満である場合、即ち、タイトル及びパフォーマ名の両方を用いた場合における演算値が閾値 X を越えたタイトル作成情報及びパフォーマ名作成情報に対応する曲の数が定数 N 個未満であるときには、情報作成サーバ 7 が候補となる曲のリスト等である曲コンテンツ選択用情報を作成するため、処理は、ステップ S

48へ進む。

【0173】

一方、タイトル及びパフォーマ名の両方を用いた場合における演算値が閾値Xを越えたタイトル作成情報及びパフォーマ名作成情報に対応する曲の数が定数N個以上のときには、情報作成サーバ7は、候補となる曲の数を絞り込むために、ユーザに曲のジャンルを入力するように要求するジャンル入力要求情報をクライアント2に送信する。制御部2eは、情報作成サーバ7から通信部2dを介して受信したジャンル入力要求情報を、出力部2bのスピーカ等に音声として出力させる。そして、処理は、ステップS47へ進む。

【0174】

続いて、ステップS47において、ユーザがジャンル入力要求情報をスピーカ等から聞いた後に、制御部2eは、ユーザが入力部2aのマイク等に音声で入力したジャンル音声情報を、この入力部2aから供給される。制御部2eは、入力部2aから供給されたジャンル音声情報を通信部2dに供給する。通信部2dは、供給されたジャンル音声情報を基地局4を介してサーバシステム3の音声認識サーバ5に送信する。

【0175】

音声認識サーバ5は、通信部2dから基地局4を介して受信したジャンル音声情報について音声認識を行う。音声認識サーバ5は、この音声認識を行った結果である認識ジャンル音声情報を情報作成サーバ7に供給する。

【0176】

情報作成サーバ7は、音響的な特徴量における、認識ジャンル音声情報と、作成情報記憶部9bに記憶されている情報提供システムで提供可能な曲のジャンルに関する1又は2以上の各ジャンル作成情報との類似性を、図示しない作成情報演算処理部を用いて演算する。この情報作成サーバ7は、作成情報記憶部9bに記憶されている全ての各ジャンル作成情報に対して、音響的な特徴量における認識ジャンル音声情報との類似性を演算する。情報作成サーバ7は、この演算値が閾値Xを越えたものを、ユーザがサーバシステム3から提供されることを希望する曲の候補になると判断する。

【0177】

そして、演算値が閾値Xを越えたジャンル作成情報に対応する曲の数が定数N個未満である場合、即ち、タイトル、パフォーマ名及びジャンルを用いた場合における演算値が閾値Xを越えたタイトル作成情報、パフォーマ名作成情報及びジャンル作成情報に対応する曲の数が定数N個未満であるときには、情報作成サーバ7が候補となる曲のリスト等である曲コンテンツ選択用情報を作成するため、処理は、ステップS48へ進む。

【0178】

一方、タイトル、パフォーマ名及びジャンルを用いた場合における演算値が閾値Xを越えたタイトル作成情報、パフォーマ名作成情報及びジャンル作成情報に対応する曲の数が定数N個以上のときには、情報作成サーバ7が候補となる曲のリスト等である曲コンテンツ選択用情報を作成するため、処理は、ステップS48へ進む。但し、この場合の曲のリスト等である曲コンテンツ選択用情報は、演算値がジャンル音声情報に最も類似していたものから順にN個の項目からなるものである。

【0179】

続いて、ステップS48において、演算値が閾値Xを越えた上記作成情報に対応する曲の数が定数N個未満である場合には、情報作成サーバ7は、N個未満の候補となる曲のリスト等である曲コンテンツ選択用情報を作成する。一方、演算値が閾値Xを越えた上記作成情報に対応する曲の数が定数N個以上の場合には、情報作成サーバ7は、演算値がジャンル音声情報に最も類似していたものから順番に取得したN個の候補となる曲のリスト等である曲コンテンツ選択用情報を作成する。

【0180】

続いて、ステップS49において、情報作成サーバ7は、作成した曲コンテンツ選択用情報を、テキスト表示等として文字で出力した方がユーザが曲を選択しやすいか否かを図示しない最適出力状態判断部を用いて判断し、文字で出力するのが最適だと判断しなかったときは、処理は、ステップS51へ進む。

【0181】

一方、情報作成サーバ7は、作成した曲コンテンツ選択用情報を、テキスト表示等として文字で出力した方がユーザが曲を選択しやすいか否かを図示しない最適出力状態判断部を用いて判断し、文字で出力するのが最適だと判断したときは、文字で出力するための曲コンテンツ選択用情報をクライアント2に送信する。

【0182】

続いて、ステップS50において、演算値が閾値Xを越えた上記作成情報に対応する曲の数が定数N個未満である場合に、情報作成サーバ7が曲コンテンツ選択用情報をクライアント2に送信したときには、制御部2eは、情報作成サーバ7から通信部2dを介して受信したN個未満の曲コンテンツ選択用情報を、出力部2bの液晶ディスプレイ等に文字として表示させる。

【0183】

一方、演算値が閾値Xを越えた上記作成情報に対応する曲の数が定数N個以上の場合に、情報作成サーバ7が曲コンテンツ選択用情報をクライアント2に送信したときには、制御部2eは、情報作成サーバ7から通信部2dを介して受信した演算値がジャンル音声情報に最も類似していたものからの順番であるN個の曲コンテンツ選択用情報を、出力部2bの液晶ディスプレイ等に文字として出力させる。

【0184】

続いて、ステップS51において、情報作成サーバ7が、作成した曲コンテンツ選択用情報を、静止画で出力した方がユーザが曲を選択しやすいか否かを図示しない最適出力状態判断部を用いて判断し、静止画で出力するのが最適だと判断しなかったときは、処理は、ステップS53へ進む。

【0185】

一方、情報作成サーバ7は、作成した曲コンテンツ選択用情報を、CDのジャケット写真等によって静止画で出力した方がユーザが曲を選択しやすいか否かを図示しない最適出力状態判断部を用いて判断し、静止画で出力するのが最適だと判断したときは、静止画で出力するための曲コンテンツ選択用情報をクライアント2に送信する。

【0 1 8 6】

続いて、ステップ S 5 2 において、演算値が閾値 X を越えた上記作成情報に対応する曲の数が定数 N 個未満である場合に、情報作成サーバ 7 が曲コンテンツ選択用情報をクライアント 2 に送信したときには、制御部 2 e は、情報作成サーバ 7 から通信部 2 d を介して受信した N 個未満の曲コンテンツ選択用情報を、出力部 2 b の液晶ディスプレイ等に静止画として出力させる。

【0 1 8 7】

一方、演算値が閾値 X を越えた上記作成情報に対応する曲の数が定数 N 個以上の場合に、情報作成サーバ 7 が曲コンテンツ選択用情報をクライアント 2 に送信したときには、制御部 2 e は、情報作成サーバ 7 から通信部 2 d を介して受信した演算値がジャンル音声情報に最も類似していたものからの順番である N 個の曲コンテンツ選択用情報を、出力部 2 b の液晶ディスプレイ等に静止画として出力させる。

【0 1 8 8】

続いて、ステップ S 5 3 において、情報作成サーバ 7 は、作成した曲コンテンツ選択用情報を、V i d e o C l i p 等によって動画で出力した方がユーザが曲を選択しやすいか否かを図示しない最適出力状態判断部を用いて判断し、動画で出力するのが最適だと判断したときは、動画で出力するための曲コンテンツ選択用情報をクライアント 2 に送信する。

【0 1 8 9】

一方、情報作成サーバ 7 は、作成した曲コンテンツ選択用情報を、動画で出力した方がユーザが曲を選択しやすいか否かを図示しない最適出力状態判断部を用いて判断し、動画で出力するのが最適だと判断しなかったときは、処理は、ステップ S 5 3 へ進む。

【0 1 9 0】

続いて、ステップ S 5 4 において、演算値が閾値 X を越えた上記作成情報に対応する曲の数が定数 N 個未満である場合に、情報作成サーバ 7 が曲コンテンツ選択用情報をクライアント 2 に送信したときには、制御部 2 e は、情報作成サーバ 7 から通信部 2 d と制御部 2 e とを介して受信した N 個未満の曲コンテンツ選択

用情報を、出力部 2 b の液晶ディスプレイ等に動画として出力させる。

【0191】

一方、演算値が閾値 X を越えた上記作成情報に対応する曲の数が定数 N 個以上の場合に、情報作成サーバ 7 が曲コンテンツ選択用情報をクライアント 2 に送信したときには、制御部 2 e は、情報作成サーバ 7 から通信部 2 d を介して受信した演算値がジャンル音声情報に最も類似していたものからの順番である N 個の曲コンテンツ選択用情報を、出力部 2 b の液晶ディスプレイ等に動画として出力させる。

【0192】

続いて、ステップ S 5 5 において、演算値が閾値 X を越えた上記作成情報に対応する曲の数が定数 N 個未満である場合に、情報作成サーバ 7 が曲コンテンツ選択用情報をクライアント 2 に送信したときには、制御部 2 e は、情報作成サーバ 7 から通信部 2 d を介して受信した N 個未満の曲コンテンツ選択用情報を、出力部 2 b のスピーカ等に曲を音声として出力させる。

【0193】

一方、演算値が閾値 X を越えた上記作成情報に対応する曲の数が定数 N 個以上の場合に、情報作成サーバ 7 が曲コンテンツ選択用情報をクライアント 2 に送信したときには、制御部 2 e は、情報作成サーバ 7 から通信部 2 d を介して受信した演算値がジャンル音声情報に最も類似していたものからの順番である N 個の曲コンテンツ選択用情報を、出力部 2 b のスピーカ等に曲を音声として出力させる。

【0194】

続いて、ステップ S 5 6 において、ユーザが曲コンテンツ選択用情報をスピーカ等から聞いた後に、制御部 2 e は、この曲コンテンツ選択用情報である曲のリストの中からユーザが選択した曲に関する情報を入力部 2 a のマイク等に音声で入力された音声情報を、この入力部 2 a から供給される。

【0195】

具体的には、制御部 2 e は、例えば、曲コンテンツ選択用情報である曲のリストの項目に割り振られた番号を入力部 2 a に音声で入力された音声情報を、この

入力部 2 a から供給される。

【0196】

また、制御部 2 e は、例えば、曲コンテンツ選択用情報である曲の名前を入力部 2 a に音声で入力された音声情報を、この入力部 2 a から供給される。

【0197】

さらに、制御部 2 e は、例えば、出力部 2 b が曲コンテンツ選択用情報である曲のリストを順番に音声で出力している場合で、現在、音声で出力されている曲をユーザが選択したいときに、「この曲」や「OK」等という言葉が入力部 2 a に音声で入力された音声情報を、この入力部 2 a から供給される。この場合、入力部 2 a は、キー入力装置やジョグダイヤル等により構成されていてもよい。

【0198】

さらにまた、制御部 2 e は、例えば、出力部 2 b が曲コンテンツ選択用情報である曲の一部分を順番に再生している場合で、現在、再生している曲をユーザが選択したいときに、「この曲」や「OK」等という言葉が入力部 2 a に音声で入力された音声情報を、この入力部 2 a から供給される。この場合も、入力部 2 a は、キー入力装置やジョグダイヤル等により構成されていてもよい。

【0199】

さらにまた、制御部 2 e は、例えば、予めクライアント 2 がサーバシステム 3 からユーザの趣向にあった曲のリストをダウンロードしてある場合に、その曲のリストの中からユーザが選択した曲の名前が入力部 2 a に音声で入力された音声情報を、この入力部 2 a から供給される。このユーザの趣向にあった曲のリストは、ユーザが予め設定しておいてもよい。

【0200】

以上の様な操作をされることにより、曲コンテンツ選択用情報である曲のリスト等の中から、ユーザがサーバシステム 3 から提供されることを希望する曲が選択される。

【0201】

続いて、ステップ S 57 において、制御部 2 e は、入力部 2 a に入力された選択音声情報を、この入力部 2 a から供給される。制御部 2 e は、入力部 2 a から

供給された選択音声情報を通信部 2 d に供給する。通信部 2 d は、供給された選択音声情報を基地局 4 を介してサーバシステム 3 の音声認識サーバ 5 に送信する。

【0202】

音声認識サーバ 5 は、通信部 2 d から基地局 4 を介して受信した選択音声情報について音声認識を行う。音声認識サーバ 5 は、この音声認識を行った結果である認識選択音声情報を、クライアント 2 に送信する。

【0203】

制御部 2 e は、音声認識サーバ 5 から通信部 2 d を介して受信した認識選択音声情報を、出力部 2 b のスピーカ等に音声として出力させる。ユーザは、この認識選択音声情報をスピーカ等から聞くことにより、自分が選択した曲をサーバシステム 3 が正しく認識したかどうか確認する。

【0204】

続いて、ステップ S 5 8 において、情報提供サーバ 8 は、音声認識サーバ 5 から供給された認識選択音声情報に基づき、クライアント 2 に送信する音楽等のコンテンツをコンテンツ情報記憶部 9 c から取得してクライアント 2 に送信する。

【0205】

制御部 2 e は、情報提供サーバ 8 から通信部 2 d を介して受信した音楽等のコンテンツを、出力部 2 b のスピーカ等に再生させる。

【0206】

続いて、ステップ S 5 9 において、出力部 2 b からのコンテンツの再生が終了し、ユーザが引き続き MOD サービスを利用する場合には、処理は、ステップ S 4 3 へ戻る。

【0207】

一方、ユーザが MOD サービスの利用を終了するためのサービス終了要求情報を、入力部 2 a のマイク等に音声で入力した場合には、制御部 2 e は、入力部 2 a に入力された音声情報であるサービス終了要求音声情報を、この入力部 2 a から供給される。制御部 2 e は、入力部 2 a から供給されたサービス終了要求音声情報を通信部 2 d に供給する。通信部 2 d は、供給されたサービス終了要求音声

情報を基地局 4 を介してサーバシステム 3 の音声認識サーバ 5 に送信する。

【0208】

音声認識サーバ 5 は、通信部 2 d から基地局 4 を介して受信したサービス終了要求音声情報について、図示しない音響分析部、音響モデル及び言語モデル等に基づき音声認識を行う。音声認識サーバ 5 は、この音声認識を行った結果である認識サービス終了要求音声情報を情報作成サーバ 7 に供給する。

【0209】

情報作成サーバ 7 は、音声認識サーバ 5 から認識サービス終了要求音声情報が供給されると、サービス終了許可情報をクライアント 2 に送信する。

【0210】

制御部 2 e は、情報作成サーバ 7 から通信部 2 d を介して受信したサービス終了許可情報を、制御部 2 e のスピーカ等に音声として出力させる。

【0211】

ユーザは、このサービス終了許可情報をスピーカ等から音声として聞くことによって、MOD サービスが終了したことを認識する。

【0212】

以上のように処理されることにより、どのように曲コンテンツ選択用情報である曲のリストが出力されればユーザが容易に目的とする曲を選択できるのかをサーバシステム 3 側が判断するため、上述した第 1 の実施の形態におけるコンテンツ選択システムよりも、ユーザは、曲コンテンツ選択用情報である曲のリストから目的とする曲をより簡単に、より速く選択することができる。

【0213】

以上述べたように、本発明を適用した実施の形態であるコンテンツ選択システム 1 では、ユーザは、入力部 2 a のマイク等に音声情報を入力することにより、出力部 2 b に出力された曲のリストの中から簡単に目的とする曲を選択することができる。従って、ユーザは、音声認識装置を備えたクライアント 2 又はサーバシステム 3 を用いることにより、簡単な操作で MOD サービスを利用することができる。

【0214】

なお、上述したコンテンツ選択システム1は、IMT-2000を用いてネットワークを構築しているが、他の通信方式のネットワークを用いてもよい。また、上述したクライアント2は、図1では便宜上1つしか示していないが、実際には複数のクライアント2がIMT-2000を介してサーバシステム3と接続されることになる。さらに、上述した音声認識機能は、音声認識サーバ5に備えられているが、クライアント2等に備えられていてもよい。

【0215】

【発明の効果】

以上説明したように、本発明に係るコンテンツ選択システム、コンテンツ選択クライアント、コンテンツ選択サーバ及びコンテンツ選択方法によれば、操作者は、クライアントの入力情報送信手段に音声情報を入力することにより、出力手段に出力されたコンテンツの中から速く簡単に目的とするコンテンツを選択することができる。

【図面の簡単な説明】

【図1】

本発明を適用した実施の形態を示すコンテンツ選択システムの構成図である。

【図2】

本発明を適用した第1の実施の形態における、MODサービスが開始されてから終了するまでの一連の処理の流れを示すフローチャートである。

【図3】

本発明を適用した第2の実施の形態における、MODサービスが開始されてから終了するまでの一連の処理の流れを示すフローチャートである。

【図4】

本発明を適用した第3の実施の形態における、MODサービスが開始されてから終了するまでの一連の処理の流れを示すフローチャートである。

【符号の説明】

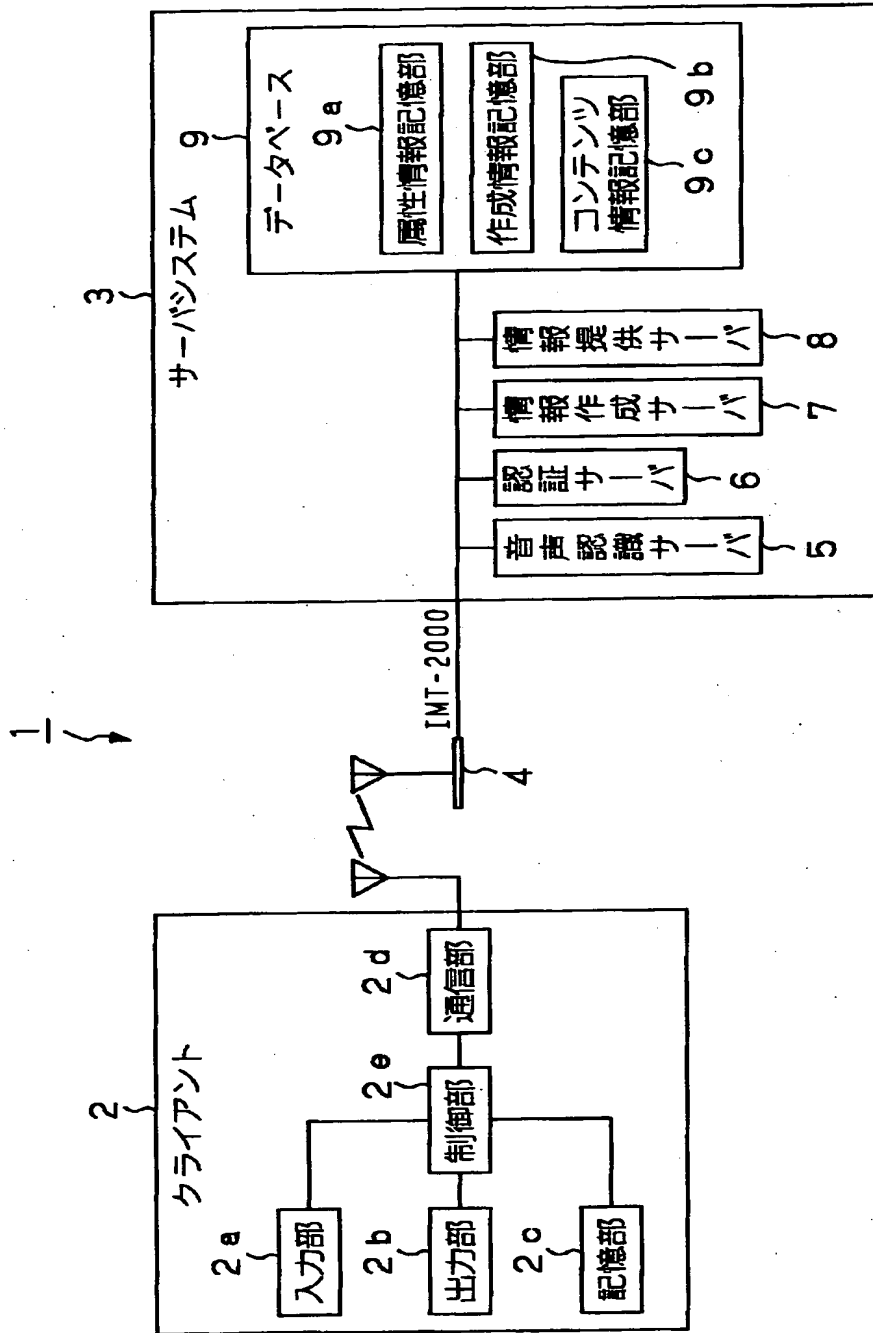
1 コンテンツ選択システム、2 クライアント、3 サーバシステム、4 基地局、5 音声認識サーバ、6 認証サーバ、7 選択サーバ、8 情報提供

特平 11-056107

サーバ、9 データベース

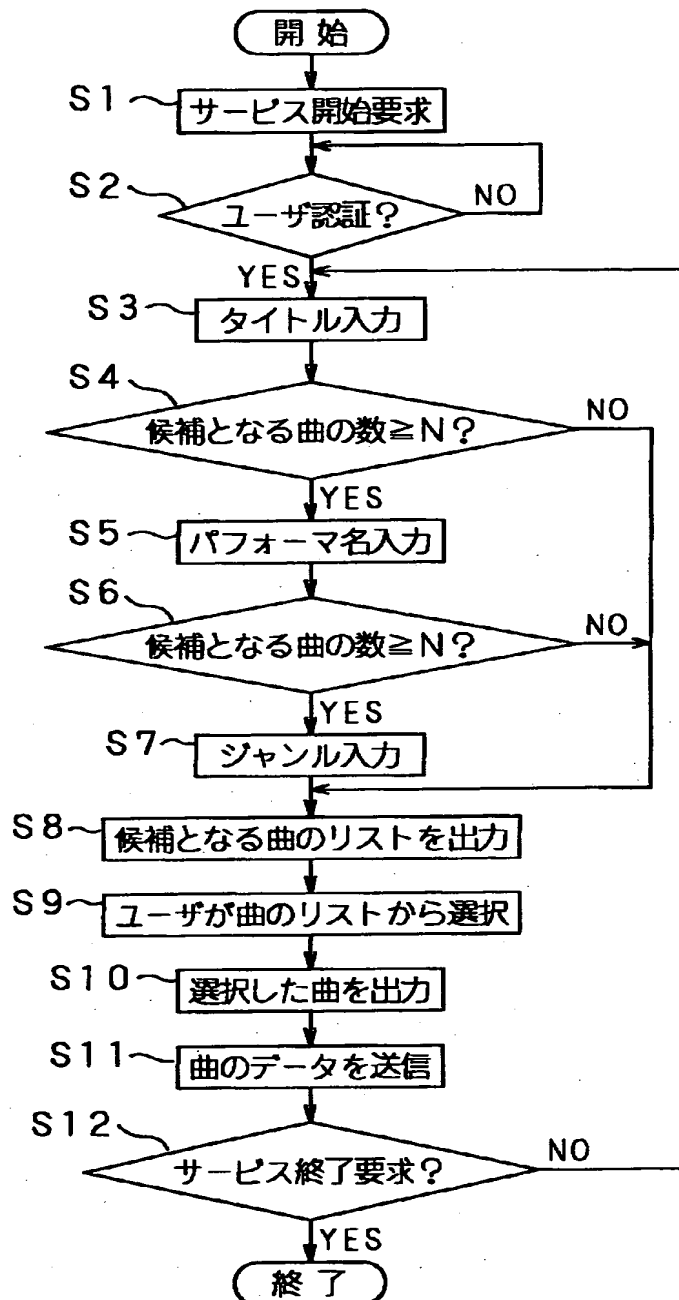
【書類名】 図面

【図 1】



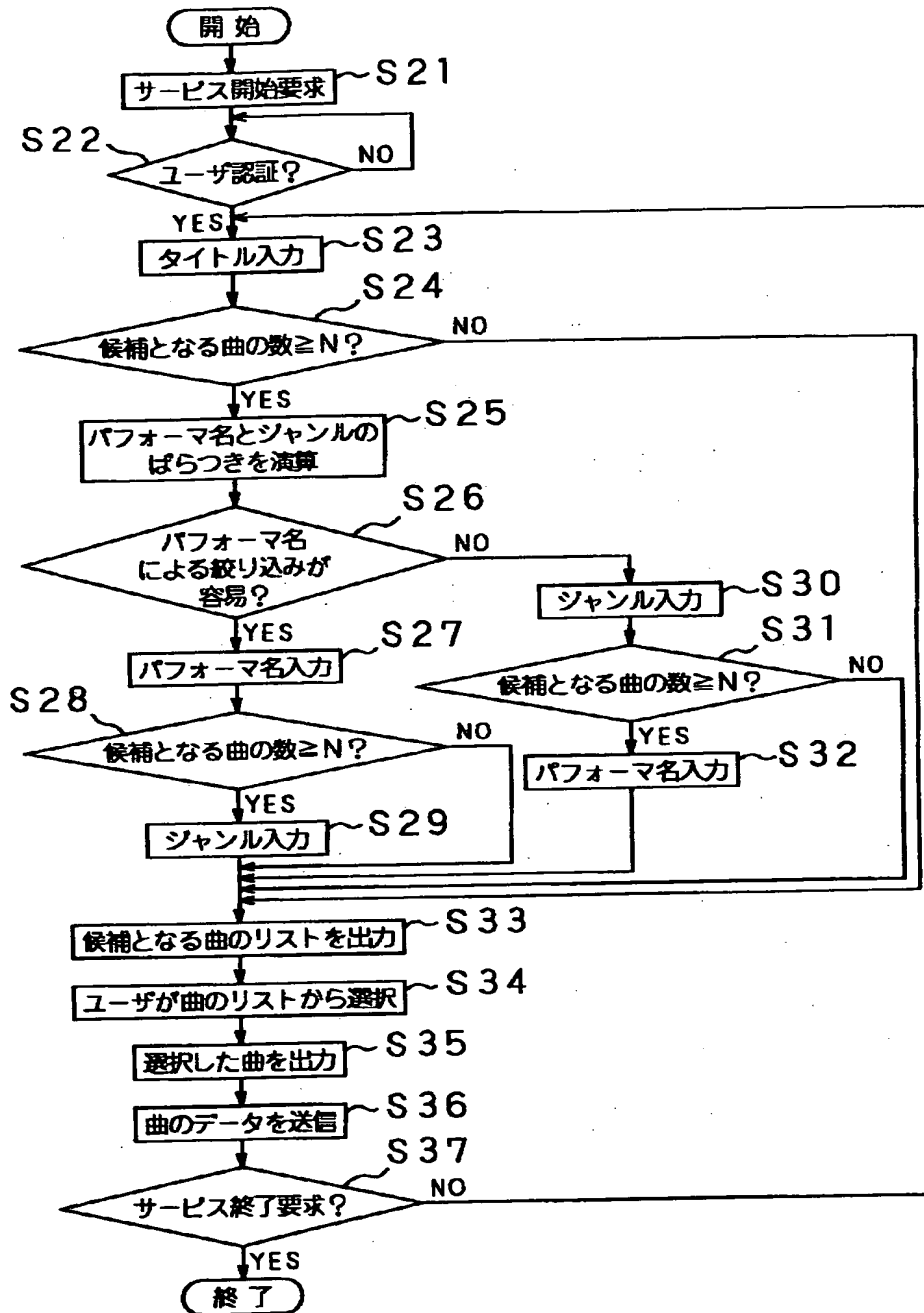
コンテンツ選択システム

【図 2】



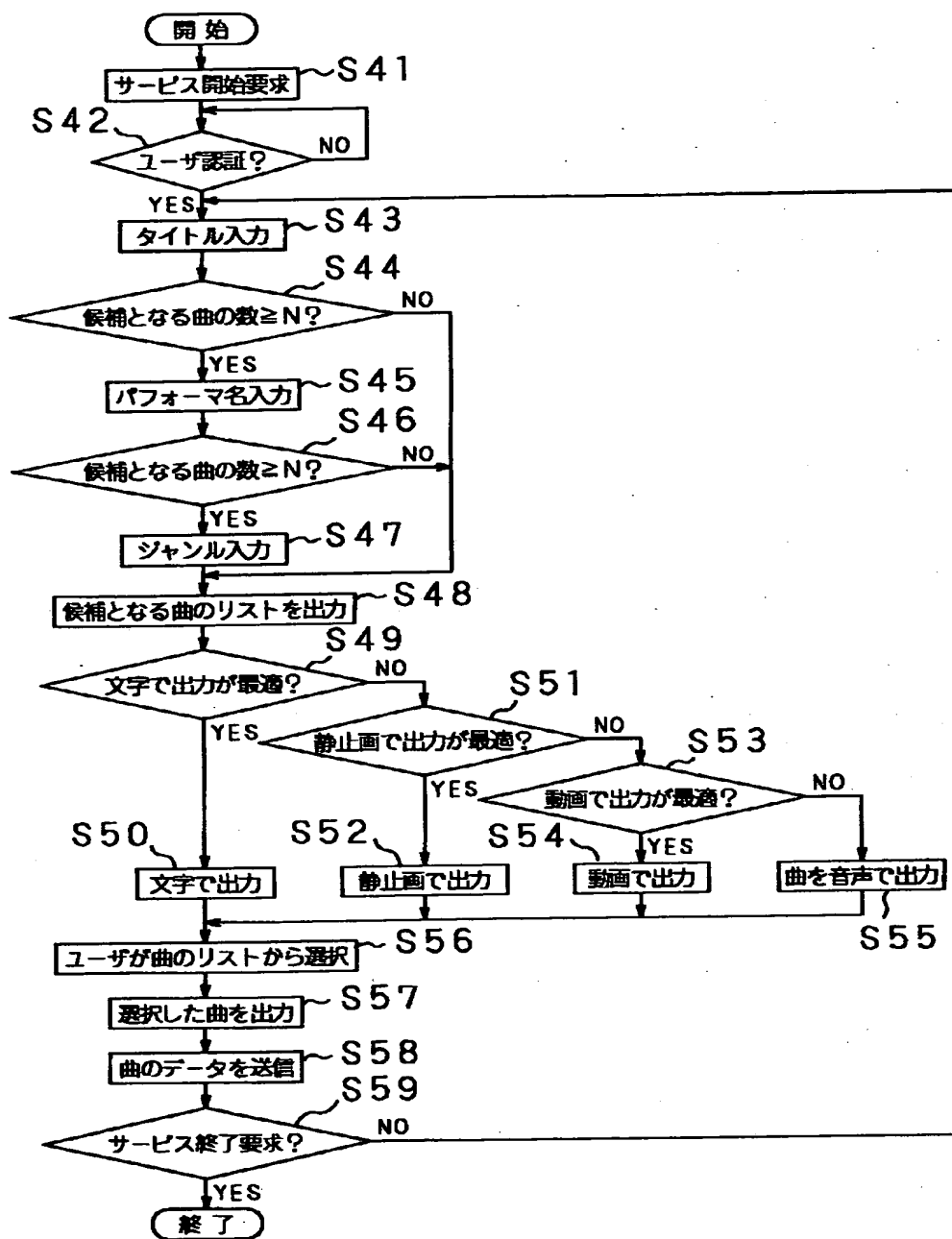
第1の実施の形態におけるフローチャート

【図 3】



第2の実施の形態におけるフローチャート

【図 4】



第 3 の実施の形態におけるフローチャート

【書類名】 要約書

【要約】

【課題】 クライアントのマイク等に音声情報を入力することにより、サーバからネットワークを介してクライアントが受信して出力した複数の項目であるコンテンツの中からどのコンテンツを取得したいのかを操作者に選択させるようにする。

【解決手段】 入力された音声情報をネットワークを介してサーバシステム3に送信する通信部2dと、サーバシステム3からネットワークを介してコンテンツ選択用情報を受信して出力する出力部2bとを備えるクライアント2と、各コンテンツに関する1又は2以上の作成情報を各コンテンツごとに記憶する作成情報記憶部9bと、クライアント2からネットワークを介して受信した音声情報と作成情報とに基づいてコンテンツ選択用情報を作成し、この作成されたコンテンツ選択用情報をクライアント2にネットワークを介して送信する情報作成サーバ7とを備えるサーバシステム3とを用いる。

【選択図】 図1

特平11-056107

出 願 人 履 歴 情 報

識別番号

[000002185]

1. 変更年月日	1990年 8月30日
[変更理由]	新規登録
住 所	東京都品川区北品川6丁目7番35号
氏 名	ソニー株式会社